



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 133468

TO: Phuong Bui  
Location: REM/2A15/2C18  
Art Unit: 1638  
Monday, September 27, 2004

Case Serial Number: 10/624061

From: Deirdre Arnold  
Location: Biotech-Chem Library  
REM 1A64  
Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

### Search Notes

## RUSH

Results are for SN 10/624,061.

*Please feel free to contact me if you have any questions or would like to amend the search.*

Thank you for using STIC services.

Regards,  
Deirdre Arnold

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2004, 02:32:10 ; Search time 68 Seconds  
(without alignments)

1211.033 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKABQSPDSGNVYTLASV.....NWVPLTKVQVYLLRLRD 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

- 1: sp\_archea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	818	61.1	265	10 Q9S7H4	Q9S7H4 arabidopsis
2	691	51.6	255	10 Q9STB2	Q9STB2 lycopersico
3	685.5	51.2	294	10 Q93Y60	Q93Y60 oryza sativ
4	689.5	50.0	316	10 Q9C544	Q9C544 arabidopsis
5	654.5	48.9	316	10 Q9XF60	Q9XF60 arabidopsis
6	432	32.3	295	3 Q96VZ8	Q96VZ8 cryptococcu
7	423	31.6	280	3 Q9P4D8	Q9P4D8 pichia angu
8	414	30.9	267	3 Q9Y7B2	Q9Y7B2 emericella
9	329	24.6	131	10 Q65861	Q65861 prunus aviu
10	283	21.1	82	10 Q22410	Q22410 petroselinu
11	188	14.0	70	10 Q22409	Q22409 petroselinu
12	110	8.2	4057	5 Q81JE7	Q81JE7 plasmodium
13	107	8.0	355	2 Q9AHC6	Q9AHC6 streptococc
14	105	7.8	360	17 Q97SJ4	Q97SJ4 streptococc
15	98.5	7.4	1259	17 Q8PU05	Q8PU05 methanosarc
16	98.5	7.4	1506	16 Q8EX35	Q8EX35 shewanella

17	96.5	7.2	668	10 Q8LPF1	Q8LPF1 arabidopsis
18	96	7.2	459	16 Q8CNG6	Q8CNG6 staphylococ
19	95.5	7.1	565	16 Q8DLU6	Q8DLU6 synchococ
20	95.5	7.1	587	16 Q99RK3	Q99RK3 staphylococ
21	95.5	7.1	587	16 Q8NV02	Q8NV02 staphylococ
22	95.5	7.1	1556	3 Q08554	Q08554 saccharomyc
23	95	7.1	281	16 Q8A4M0	Q8A4M0 bacteroides
24	95	7.1	746	11 Q8C9T7	Q8C9T7 mus musculi
25	94.5	7.1	761	16 Q89AS4	Q89AS4 buchnera ap
26	94	7.0	1561	4 Q9ULG1	Q9ULG1 homo sapien
27	93.5	7.0	1134	17 Q9UY58	Q9UY58 pyrococcus
28	93	6.9	422	16 Q8XNP6	Q8XNP6 clostridium
29	93	6.9	958	10 Q8GUL2	Q8GUL2 arabidopsis
30	92	6.9	296	3 Q9UVD5	Q9UVD5 nilaparvata
31	92	6.9	371	16 Q9ZDP9	Q9ZDP9 rickettsia
32	92	6.9	812	16 Q8G0I7	Q8G0I7 brucella su
33	92	6.9	823	16 Q8YHC6	Q8YHC6 brucella me
34	92	6.9	1844	5 Q8IE31	Q8IE31 plasmodium
35	91.5	6.8	581	9 Q8SBQ2	Q8SBQ2 bacterioph
36	90.5	6.8	333	17 Q8U3T5	Q8U3T5 pyrococcus
37	90.5	6.8	792	13 Q90ZB6	Q90ZB6 xenopus lae
38	90	6.7	274	2 Q47742	Q47742 enterococcu
39	90	6.7	660	5 Q8IIX6	Q8IIX6 plasmodium
40	90	6.7	1015	5 Q94705	Q94705 physarum po
41	89.5	6.7	251	16 Q51460	Q51460 borrelia bu
42	89.5	6.7	405	16 Q8DTJ3	Q8DTJ3 streptococ
43	89.5	6.7	534	2 Q9ZFS4	Q9ZFS4 bacteroides
44	89.5	6.7	1176	16 Q8RDW3	Q8RDW3 fusobacteri
45	89.5	6.7	1187	10 Q9ZV43	Q9ZV43 arabidopsis

#### ALIGNMENTS

RESULT 1

Q9S7H4	ID	Q9S7H4	PRELIMINARY;	PRT;	265 AA.
AC	Q9S7H4;	AC	Q9S7H4;		
DT	01-MAY-2000	(TREMELrel. 13, Created)			
DT	01-MAY-2000	(TREMELrel. 13, Last sequence update)			
DT	01-OCT-2003	(TREMELrel. 25, Last annotation update)			
DE	Chorismate mutase (EC 5.4.99.5)	(Putative chorismate mutase CM2).			
GN	CM2 OR T3ON20.140 OR ATSG10870.				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi				
OX	NCBI_TaxID=3702;				
OX	[1]				
RP	SEQUENCE FROM N.A.				
RA	Kuhn R., Vogt E., Schmid J., Amrhein N., Schaller A.;				
RT	"Expression analysis of Arabidopsis thaliana genes for plastidic (CM1)				
RT	and cytosolic (CM2) chorismate mutases."				
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=97111372; PubMed=8953244;				
RA	Eberhard J., Ehrler T.T., Eppe P., Felix G., Raesecke H.R.,				
RA	Amrhein N., Schmid J.;				
RT	"Cytosolic and plastidic chorismate mutase isozymes from Arabidopsis				
RT	thaliana: molecular characterization and enzymatic properties."				
RL	Plant J. 10:815-821(1996).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RA	Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,				
RA	Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;				
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Arabidopsis sequencing project;				
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE FROM N.A.				

Yamada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K., Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;  
 RA "Full Length cDNA of gene At5g10870 (GI:15233285).";  
 RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W., Lee J.M., Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C., Wu H.C., Yu G., Yuan S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R., Theologis A.;  
 RA "Arabidopsis Open Reading Frame (ORF) Clones";  
 RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AJ242648; CAB54519.1; -  
 DR EMBL; L47355; AAD48922.1; -  
 DR EMBL; AL365234; CAB96842.1; -  
 DR EMBL; AY065238; AAL38714.1; -  
 DR EMBL; AV133840; AAM91774.1; -  
 DR HSP; P32178; 2CSM.  
 DR GO; GO:0004106; F:chorismate mutase activity; IEA.  
 DR GO; GO:0016853; P:isomerase activity; IEA.  
 DR GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.  
 DR InterPro; IPR002701; Chorisamat mut.  
 DR InterPro; IPR008951; Chorisamat mut II.  
 DR InterPro; IPR008238; Chor mut\_AroQ\_eu.  
 DR Pfam; PF01817; Chorisamat mut; 1.  
 DR PIRSP; PIRSF017318; Chor\_mut\_AroQ\_eu; 1.  
 KW isomerase.  
 SQ SEQUENCE 265 AA; 30491 MW; 8F2616AE7482FA00 CRC64;  
 Query Match 61.1%; Score 818; DB 10; Length 265;  
 Best Local Similarity 61.1%; Pred. No. 2.7e-50;  
 Matches 162; Conservative 31; Mismatches 68; Indels 4; Gaps 2;  
 QY 1 MAKAAQSPDSG--NYTTLASVREDIVROEDTIYGLIERAKFPNSHTYDKYAIQGF 58  
 DB 1 MARVFSDSGSGSNVLSLDLRESLRQEDTIVFSLIERAKFPNSPAFESRCLDSGS 60  
 QY 59 CGSLVFEVKNATIAQAKAGRYKNPENAFPPENLPPSIVPSYFKQFLHPCGAASININK 118  
 DB 61 FSSLTEFFVRETEIIGAKVGRYEYEPENPFENIPHSVFTPKYPSALHPKALSVNINK 120  
 QY 119 SIWKYFKELLPLLATSGDGNYAQTAANDLSLQSISSRIHYGKFAVVKFRDAPQDY 178  
 DB 121 QIWDIYFKELLPLLVKPGDGNYPSTAAASDLACLQALSRIHYGKFAVVKFRDAPQDY 180  
 QY 179 PLTRAKDKGLMKLLFTSVETVRKVEKAVVFGQVNLNSDDNENRKF--DPSVA 236  
 DB 181 PAIRAQDREALMKLLTFEYKVEEMVKRQVKAETFGQEVKFNSGYGDGSKKIKVDPLIA 240  
 QY 237 SSIYKNWVPLTKVEOVYLLRLD 261  
 DB 241 SRIYGEWLPLTKLVEVYLLRLD 265  
 RESULT 2  
 Q9STB2 PRELIMINARY; PRT; 255 AA.  
 AC Q9STB2  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Chorismate mutase (EC 5.4.99.5)  
 OS Lycopersicon esculentum (Tomato)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.  
 NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UC82b;  
 RX MEDLINE=96400046; PubMed=8806422;  
 RA Eberhard J., Bischoff M., Raesecke H.R., Amrhein N., Schmid J.;  
 RT "Isolation of a cDNA from tomato coding for an unregulated, cytosolic chorismate mutase";  
 RL Plant Mol. Biol. 31:917-922(1996).  
 DR EMBL; L47356; AAD48923.1; -  
 DR HSP; P32178; 2CSM.  
 DR GO; GO:0004106; F:chorismate mutase activity; IEA.  
 DR GO; GO:0016853; P:isomerase activity; IEA.  
 DR GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.  
 DR InterPro; IPR002701; Chorisamat mut.  
 DR InterPro; IPR008951; Chorisamat mut II.  
 DR InterPro; IPR008238; Chor mut\_AroQ\_eu.  
 DR Pfam; PF01817; Chorisamat mut; 1.  
 DR PIRSP; PIRSF017318; Chor\_mut\_AroQ\_eu; 1.  
 KW isomerase.  
 SQ SEQUENCE 255 AA; 29215 MW; C32B54975A578B9A CRC64;  
 Query Match 51.6%; Score 691; DB 10; Length 255;  
 Best Local Similarity 55.7%; Pred. No. 1.1e-49;  
 Matches 137; Conservative 39; Mismatches 68; Indels 2; Gaps 2;  
 QY 17 LASVREDIVROEDTIYGLIERAKFPNSHTYDKYAIQGFQCSLVFVKNTAIOAK 76  
 DB 11 LDSIRKSLIRQEDTIIFNLIERIKFPINFTLYKQQLPPSSNFGSLFQVLFQETSLQSK 70  
 QY 77 AGRYKNPENAFPPENLPPSIVPSYFKQFLHPCGAASININKSIWKYFKELLPLLAT- 135  
 DB 71 VGRYLAPEENPFPPDNLSLSIIPLTKTPVLHPAESVNVNEKILDIYNQMLPLFCTEV 130  
 QY 136 GDGNYAQTAANDLSLQSISSRIHYGKFAVVKFRDAPQDYPLIRAKDKGLMKLLTF 195  
 DB 131 NDDANFATTAACTIQLLQALSRIHYGKFAVVKFRDSDIYKPFILAQDRDALMKLLTF 190  
 QY 196 TSVEETVRKVEKAVVFGQVNLNSDDNENRKFDPSPVASSLYKNWVPLTKVEQVEY 255  
 DB 191 EAVEEMVKKVAKAKVFGQVSLN-DNAEVEKGIKIDPLLVSLYDEWVWPLTKLVEVEY 249  
 QY 256 LLRLD 261  
 DB 250 LLRLD 255  
 RESULT 3  
 Q93Y60 PRELIMINARY; PRT; 294 AA.  
 ID Q93Y60  
 AC Q93Y60  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative chorismate mutase.  
 GN P04030C05.9  
 OS Oryza sativa (Rice)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone:P04030C05";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003239; BAB63589.1; -  
 DR Gramene; Q93Y60; -  
 DR GO; GO:0004106; F:chorismate mutase activity; IEA.



GO; GO:0009073; P:aromatic amino acid family biosynthesis; IEA.

RC  
RA Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,  
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Heuan V.W.,  
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
RA Pang C.C., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
RA Tang C.J., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,  
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
RA "Arabidopsis ORF clones";  
RP Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

Query Match 51.2%; Score 685.5; DB 10; Length 294;  
Best Local Similarity 51.5%; Pred. No. 3.7e-49;  
Matches 135; Conservative 45; Mismatches 81; Indels 1; Gaps 1;

QY 1 MAKAAEQSPDGNVYTLASVREDLVROEDTIYGLIERAKFSPNSHTYDEKVAQIOGFCG 60  
DB 33 VYKEEKQRIQDSEITLDIRTSVLRQDSIFSLERAAQFCYNADIYDKNAHVDGFDG 92  
QY 61 SLVEFVVKNTAIAQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININKSI 120  
DB 93 SLVEFVVRTEKHLHQGVGRKSPDHPFPFEDLPPELPLPQYKVLHPIADSNINKEI 152  
QY 121 WKMYFKEPLPLATSGDDQNYAQTAAANDLSLQSISSRIHYGKFAEVRKFRDAPQDYEP 180  
DB 153 WKMYFDELLPRVLKSGSDGNYGSSALCDTICLQALSKRIHYGKFAEVRKFRDAPQDYEP 212  
QY 181 IRAKQEGKMLKLTTSVEETVRKVEKAVVFGQEVNLSNDNDNENR-KFQPSVASSL 239  
DB 213 IIAQDCQDLMLLTVEYTERAIEHRVAKAKIFGQEVNLSNDNDNENR-KFQPSVASSL 272  
QY 240 YKNWIPILTKVQVYLLRLD 261  
DB 273 YSTRIMPLTKVQVYLLRLD 294

RESULT 4  
Q9C544 PRELIMINARY; PRT; 316 AA.

ID Q9C544  
AC Q9C544; 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Chorisimate mutase, putative (Putative chorisimate mutase, 16810-15349)  
DE (Atlg69370).  
GN F10D13.6 OR F3010.5 OR ATG69370/F23010.5.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.

STRAIN=cv. Columbia;  
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,  
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,  
RA Hayashizaki Y., Shinozaki K.;  
RA "Arabidopsis thaliana full-length cDNA";  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

QY 10 DSGNVYTLASVREDLVROEDTIYGLIERAKFSPNSHTYDEKVAQIOGFCGSLVEFVVK 69  
DB 64 DESEYKLESIRHSLIRQDSIFLNLERAQRYNADTYDEDAFTMEGFGSLVEFVVK 123  
QY 70 TEATQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININKSIKWYFKELL 129  
DB 124 TEKHAHVDRYKSPDEHPFPQCLPELIPPIQYQVQLHRCASININKKVNMYFKHLL 183  
QY 130 PLLATSGDDGNYAQTAAANDLSLQSISSRIHYGKFAEVRKFRDAPQDYEPILIRAKQEG 189  
DB 184 PRLVKPGDDGNGCGSAALCDTMCQLQILSKRIHFQKFAEVRKFRDAPQDYEPILIRAKQ 243  
QY 190 MKLLTFTSVEETVRKVEKAVVFGQEVNLSNDNDNENR-KFQPSVASSLYKNWIPIL 248  
DB 244 MQLTYETVEVVRKVEIKARIFGQDITINDPETADEPSYKIQPSLVAKLYGERIMPLT 303

Query Match 50.0%; Score 669.5; DB 10; Length 316;  
Best Local Similarity 51.4%; Pred. No. 8.8e-48;  
Matches 130; Conservative 44; Mismatches 78; Indels 1; Gaps 1;

QY 10 DSGNVYTLASVREDLVROEDTIYGLIERAKFSPNSHTYDEKVAQIOGFCGSLVEFVVK 69  
DB 64 DESEYKLESIRHSLIRQDSIFLNLERAQRYNADTYDEDAFTMEGFGSLVEFVVK 123  
QY 70 TEATQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININKSIKWYFKELL 129  
DB 124 TEKHAHVDRYKSPDEHPFPQCLPELIPPIQYQVQLHRCASININKKVNMYFKHLL 183  
QY 130 PLLATSGDDGNYAQTAAANDLSLQSISSRIHYGKFAEVRKFRDAPQDYEPILIRAKQEG 189  
DB 184 PRLVKPGDDGNGCGSAALCDTMCQLQILSKRIHFQKFAEVRKFRDAPQDYEPILIRAKQ 243  
QY 190 MKLLTFTSVEETVRKVEKAVVFGQEVNLSNDNDNENR-KFQPSVASSLYKNWIPIL 248  
DB 244 MQLTYETVEVVRKVEIKARIFGQDITINDPETADEPSYKIQPSLVAKLYGERIMPLT 303

Query Match 51.4%; Score 669.5; DB 10; Length 316;  
Best Local Similarity 51.4%; Pred. No. 8.8e-48;  
Matches 130; Conservative 44; Mismatches 78; Indels 1; Gaps 1;

QY 10 DSGNVYTLASVREDLVROEDTIYGLIERAKFSPNSHTYDEKVAQIOGFCGSLVEFVVK 69  
DB 64 DESEYKLESIRHSLIRQDSIFLNLERAQRYNADTYDEDAFTMEGFGSLVEFVVK 123  
QY 70 TEATQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININKSIKWYFKELL 129  
DB 124 TEKHAHVDRYKSPDEHPFPQCLPELIPPIQYQVQLHRCASININKKVNMYFKHLL 183  
QY 130 PLLATSGDDGNYAQTAAANDLSLQSISSRIHYGKFAEVRKFRDAPQDYEPILIRAKQEG 189  
DB 184 PRLVKPGDDGNGCGSAALCDTMCQLQILSKRIHFQKFAEVRKFRDAPQDYEPILIRAKQ 243  
QY 190 MKLLTFTSVEETVRKVEKAVVFGQEVNLSNDNDNENR-KFQPSVASSLYKNWIPIL 248  
DB 244 MQLTYETVEVVRKVEIKARIFGQDITINDPETADEPSYKIQPSLVAKLYGERIMPLT 303

Query Match 50.0%; Score 669.5; DB 10; Length 316;  
Best Local Similarity 51.4%; Pred. No. 8.8e-48;  
Matches 130; Conservative 44; Mismatches 78; Indels 1; Gaps 1;

QY 10 DSGNVYTLASVREDLVROEDTIYGLIERAKFSPNSHTYDEKVAQIOGFCGSLVEFVVK 69  
DB 64 DESEYKLESIRHSLIRQDSIFLNLERAQRYNADTYDEDAFTMEGFGSLVEFVVK 123  
QY 70 TEATQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININKSIKWYFKELL 129  
DB 124 TEKHAHVDRYKSPDEHPFPQCLPELIPPIQYQVQLHRCASININKKVNMYFKHLL 183  
QY 130 PLLATSGDDGNYAQTAAANDLSLQSISSRIHYGKFAEVRKFRDAPQDYEPILIRAKQEG 189  
DB 184 PRLVKPGDDGNGCGSAALCDTMCQLQILSKRIHFQKFAEVRKFRDAPQDYEPILIRAKQ 243  
QY 190 MKLLTFTSVEETVRKVEKAVVFGQEVNLSNDNDNENR-KFQPSVASSLYKNWIPIL 248  
DB 244 MQLTYETVEVVRKVEIKARIFGQDITINDPETADEPSYKIQPSLVAKLYGERIMPLT 303

Query Match 50.0%; Score 669.5; DB 10; Length 316;  
Best Local Similarity 51.4%; Pred. No. 8.8e-48;  
Matches 130; Conservative 44; Mismatches 78; Indels 1; Gaps 1;

QY 10 DSGNVYTLASVREDLVROEDTIYGLIERAKFSPNSHTYDEKVAQIOGFCGSLVEFVVK 69  
DB 64 DESEYKLESIRHSLIRQDSIFLNLERAQRYNADTYDEDAFTMEGFGSLVEFVVK 123  
QY 70 TEATQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININKSIKWYFKELL 129  
DB 124 TEKHAHVDRYKSPDEHPFPQCLPELIPPIQYQVQLHRCASININKKVNMYFKHLL 183  
QY 130 PLLATSGDDGNYAQTAAANDLSLQSISSRIHYGKFAEVRKFRDAPQDYEPILIRAKQEG 189  
DB 184 PRLVKPGDDGNGCGSAALCDTMCQLQILSKRIHFQKFAEVRKFRDAPQDYEPILIRAKQ 243  
QY 190 MKLLTFTSVEETVRKVEKAVVFGQEVNLSNDNDNENR-KFQPSVASSLYKNWIPIL 248  
DB 244 MQLTYETVEVVRKVEIKARIFGQDITINDPETADEPSYKIQPSLVAKLYGERIMPLT 303

Query Match 50.0%; Score 669.5; DB 10; Length 316;  
Best Local Similarity 51.4%; Pred. No. 8.8e-48;  
Matches 130; Conservative 44; Mismatches 78; Indels 1; Gaps 1;

QY 10 DSGNVYTLASVREDLVROEDTIYGLIERAKFSPNSHTYDEKVAQIOGFCGSLVEFVVK 69  
DB 64 DESEYKLESIRHSLIRQDSIFLNLERAQRYNADTYDEDAFTMEGFGSLVEFVVK 123  
QY 70 TEATQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININKSIKWYFKELL 129  
DB 124 TEKHAHVDRYKSPDEHPFPQCLPELIPPIQYQVQLHRCASININKKVNMYFKHLL 183  
QY 130 PLLATSGDDGNYAQTAAANDLSLQSISSRIHYGKFAEVRKFRDAPQDYEPILIRAKQEG 189  
DB 184 PRLVKPGDDGNGCGSAALCDTMCQLQILSKRIHFQKFAEVRKFRDAPQDYEPILIRAKQ 243  
QY 190 MKLLTFTSVEETVRKVEKAVVFGQEVNLSNDNDNENR-KFQPSVASSLYKNWIPIL 248  
DB 244 MQLTYETVEVVRKVEIKARIFGQDITINDPETADEPSYKIQPSLVAKLYGERIMPLT 303

Query Match 50.0%; Score 669.5; DB 10; Length 316;  
Best Local Similarity 51.4%; Pred. No. 8.8e-48;  
Matches 130; Conservative 44; Mismatches 78; Indels 1; Gaps 1;

QY 10 DSGNVYTLASVREDLVROEDTIYGLIERAKFSPNSHTYDEKVAQIOGFCGSLVEFVVK 69  
DB 64 DESEYKLESIRHSLIRQDSIFLNLERAQRYNADTYDEDAFTMEGFGSLVEFVVK 123  
QY 70 TEATQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININKSIKWYFKELL 129  
DB 124 TEKHAHVDRYKSPDEHPFPQCLPELIPPIQYQVQLHRCASININKKVNMYFKHLL 183  
QY 130 PLLATSGDDGNYAQTAAANDLSLQSISSRIHYGKFAEVRKFRDAPQDYEPILIRAKQEG 189  
DB 184 PRLVKPGDDGNGCGSAALCDTMCQLQILSKRIHFQKFAEVRKFR

Query Match	31.6%	Score 423;	DB 3;	Length 280;
Best Local Similarity	38.0%;	Pred. No. 3e-27;		
Matches	95;	Conservative 49;	Mismatches 102;	Indels 4; Gaps 3;
Qy	14	VYTLASVRDLVRQEDTIITVGLIERAFKFSNSHTYDEKVAIQGCGSLVEFVVKQTEAI	73	
Db	9	VLDLGNIRDALVRMEDTIITFNFIERSQFVYASPSVYKVNQFPIFNFGSFLDWLLSQHERI	68	
Qy	74	QAKAGYKQKPEENAFPPENLPPSIVPSYSGKQFLHFGAASININKSIWKMYFKELP-LL	132	
Db	69	HSQVRKYDAPDEVFFPPFVLEKTFLEPKINYPSVLASYADEINYNKIKLITYTSEIVPGIA	128	
Qy	133	ATSGD-DGNYQAFAANDLSLQISIRRIHYGKVAEVKFRDAPQDYEPITRAKDEGLMK	191	
Db	129	AGSGEQEDNLGSCAMADIECLQSLSRRIHGFGRVFAEAKFISEGDKIVDLKKRDVEGIEA	188	
Qy	192	LLTFTSVETVRKRVKKAVVQGVENLNSDDNDNENRRKDPSPVASSLYKNVITPLTKEV	251	
Db	189	LITNABVEKRILDRLEKGRAYGDTPLKF--TQHIOQSKVKEVITVKIYKDFVILTKKV	246	
Qy	252	QVEYILLRLRD	261	



```

FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9154 MW; AF5A191C43D5030B CRC64;

Query Match 21.1%; Score 283; DB 10; Length 82;
Best Local Similarity 53.4%; Pred. No. 2.9e-16;
Matches 52; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 88 PFENLPPSIVSYSKQFLHFGAASININKSIWKYFKELLPLLATSGDDGNYAQTAA 147
D 1 FPDNLPSLLPCQYHNPQILHPTAAININETIWDAYVQNQLLPTTEGDDGNYVPTATS 60
QY 148 DLSLQISRRIRHYGKFAEVK 169
D 61 DLQCLQISRRIRHYGKFAEVK 82

RESULT 11
O22409 PRELIMINARY; PRT; 70 AA.
AC O22409
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plastidic chorismate mutase 1 (Fragment).
GN PCW1.
OS Petroselinum crispum (Parsley) (Petroselinum hortense).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Apiales; Apiaceae; Apioideae; apioid superclade;
OC Apium clade; Petroselinum.
OC NCBI_TaxID=4043;
RN [1]
SEQUENCE FROM N.A.
RA Batz O.; Logemann E.; Reinold S.; Hahlbrock K.;
RT "Extensive reprogramming of cellular metabolism by fungal elicitor or
RT infection in parsley suggests a new perception of 'defense-related'
RT genes.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF012866; AAB69322.2; -.
DR PIR; T14901; T14901.
DR HSP; P32178; SC5M.
DR InterPro; IPR008951; Chorismat_mut_II.
FT NON_TER 1
FT NON_TER 70
SQ SEQUENCE 70 AA; 7917 MW; 5D2D266E7ACA3843 CRC64;

Query Match 14.0%; Score 188; DB 10; Length 70;
Best Local Similarity 51.4%; Pred. No. 2e-08;
Matches 36; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

QY '90 PENLPPSIVSYSKQFLHFGAASININKSIWKYFKELLPLLATSGDDGNYAQTAA 149
D 1 PEHLPPISPLQYQVLPVPTESINIKSIWKYFNLVPLVKEGEGNYGATAVCDT 60
QY 150 SLQISRRIR 159
D 61 ICLQALSRI 70

RESULT 12
O81UE7 PRELIMINARY; PRT; 4057 AA.
AC O81UE7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN Pf10_0251.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX STRAIN=3D7;
RA MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W., K.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Pailand A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McRadden G.I., Cummings L.M., Subramanian G.M., Mungall C., R.W.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AE014933; AAN35448.1; -.
KW Hypothetical protein.
SQ SEQUENCE 4057 AA; 488766 MW; 4E847AD8C1210BD8 CRC64;

Query Match 8.2%; Score 110; DB 5; Length 4057;
Best Local Similarity 21.2%; Pred. No. 15;
Matches 64; Conservative 63; Mismatches 91; Indels 84; Gaps 20;

QY 14 VYTLASVREDLVRQEDTIIYGLIERAKFPSN-----SHTYDEKY---AQICGCGSLVEFV 66
D 1559 VYFINNINELI-----TYIKIMDELFFIQNDKYKHINIEYKLSKQLLNFKNTLV--V 1611
QY 67 VKNTEALQAKA-GRYKNPE-----ENAFPELPPSIVPSYSF---KQFLHFGAASINI 116
D 1612 IKKIQTYYINTLQYKXNNEIKKCTNDEYF-----SSLINTYKHCEEFYVP-----QI 1661
QY 117 NKSIRKMY-----FKELLPLLATSG-----DGNVAQTAA--DLS 150
D 1662 NKHLEKNYDIQYFYNLLKQLELITVNLVILNKRKEFFSYLLNDENFLNIYLNKIDIK 1721
QY 151 -----LQISRRIRHYGKFAEVKFRDAPQ-DYEPLIRAKDKGLMKLITFTSVET 201
D 1722 TLNKYIHIIPSIKKLYIQNEICGVLSKQNDQIIIFKEQIKYKKBFLNLQINL-EIKW 1780
QY 202 VRKEVE---KKAVVFGQEVN-LNSDD---NDNENRKPDPSSVASSLYKNWVPIPTKEVQVE 254
D 1781 VLKNIQVNLQNEIIQKIDINYINDDNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN 1833
QY 255 YL 256
D 1834 FL 1835

RESULT 13
O9AHC6 PRELIMINARY; PRT; 355 AA.
ID O9AHC6
AC O9AHC6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Wc1M.
GN Wc1M.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=21116929; PubMed=11179285;
RA Jiang S.M., Wang L., Reeves P.R.;
RT "Molecular characterization of Streptococcus pneumoniae type 4, 6B, 8,
RT and 18C capsular polysaccharide gene clusters.";
RL Infect. Immun. 69:1244-1255(2001).
DR EMBL; AF316639; AAK20675.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0008168; F:methyltransferase activity; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.

```

DR	GO; GO:0008152; P:metabolism; IEA.	
DR	GO; GO:0006412; P:protein biosynthesis; IEA.	
DR	InterPro; IPR000787; Cor/por Metransf.	
DR	InterPro; IPR007345; PS_pyruv trans.	
DR	InterPro; IPR001865; Ribosomal S2.	
DR	Pfam; PF04230; PS_pyruv trans; 1.	
DR	PROSITE; PS00962; RIBOSOMAL_S2_1.	
SQ	SEQUENCE 355 AA; 40777 MW; 60BDC233A233D6E7 CRC64;	
Query Match		
Best Local Similarity 23.8%; Pred. No. 1;		
Matches 53; Conservative 36; Mismatches 76; Indels 58; Gaps 10;		
QY	31 IIVGLIERAKFPNSHT-----YDEKIAIQGFCGSLVEFVVVNTAIOAKAGRYKNP 83	
DB	21 IVDLLESKGYEASFPNPSDFQMYLKEYRKQ-----SFTKQADAILYIPGGYFGE 73	
QY	84 EENAFPPENLPPSIVPSYFQFLHPGAASININKSIWKYFKELPLLATSGDDGNVY 143	
DB	74 GHNARFRDNL-----IQKRFPLG-----IWASYFKKPIGVLGIGAGPN---- 113	
QY	144 TAANDLSLQISRIHYGKFAEYKFRD-----AP--QDYEPILIRAKDKEGLM 190	
DB	114 ---NDSLUNYGIKRIINHAQFIT--VRDRESFDSLKHLSPSPVHETFDLIISKLREKT 169	
QY	191 KLITFTSVEETVRKV-----EKKAV-VFGQEVNLSDDNDN 226	
DB	170 EQLCOLKREAKDKIILVYHNHKKALEKFAESISLFLNPN 212	
RESULT 14		
Q975J4	Q975J4	
ID	Q975J4	PRELIMINARY; PRT; 360 AA.
AC	Q975J4;	
DT	01-OCT-2001 (TrEMBLrel. 18, Created)	
DT	01-OCT-2001 (TrEMBLrel. 18, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	Hypothetical protein SP0355.	
GN	SP0355.	
OS	Streptococcus pneumoniae.	
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
OC	Streptococcus.	
OX	NCBI_TaxID=1313;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC BAA-334 / TIGR4;	
RX	MEDLINE=21357209; PubMed=11463916;	
RA	Tetelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,	
RA	Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,	
RA	Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,	
RA	Mayhew L.A., White O., Salzberg S.L., Lewis M.R., Rade D.,	
RA	Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,	
RA	McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey B.K.,	
RA	Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,	
RA	Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;	
RT	"Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.";	
RL	Science 293:498-506(2001).	
DR	EMBL; AE007347; AAK74528.1; -.	
DR	PIR; G95041; G95041.	
DR	TIGR; SP0355; -.	
DR	GO; GO:0005622; C:intracellular; IEA.	
DR	GO; GO:0005840; C:ribosome; IEA.	
DR	GO; GO:0006168; F:methyltransferase activity; IEA.	
DR	GO; GO:0003735; P:structural constituent of ribosome; IEA.	
DR	GO; GO:0008152; P:metabolism; IEA.	
DR	GO; GO:0006412; P:protein biosynthesis; IEA.	
DR	InterPro; IPR000878; Cor/por Metransf.	
DR	InterPro; IPR007345; PS_pyruv trans.	
DR	Pfam; PF04230; PS_pyruv trans; 1.	
DR	PROSITE; PS00962; RIBOSOMAL_S2_1.	
KW	Hypothetical protein; Complete proteome.	
Query Match		
Best Local Similarity 25.9%; Pred. No. 29;		
Matches 59; Conservative 36; Mismatches 86; Indels 47; Gaps 15;		
SQ	SEQUENCE 1259 AA; 145131 MW; 9F523C6BF7A78440 CRC64;	
Query Match		
Best Local Similarity 24.7%; Pred. No. 1.6;		
Matches 55; Conservative 34; Mismatches 76; Indels 58; Gaps 11;		
QY	31 IIVGLIERAKF-----PSN-SHTYDEKIAIQGFCGSLVEFVVVNTAIOAKAGRYKNP 83	
DB	26 IVDLLESKGYEVSFPNPSDFQMYLKEYRKQ-----SFTKQADAILYIPGGYFGE 78	
QY	84 EENAFPPENLPPSIVPSYFQFLHPGAASININKSIWKYFKELPLLATSGDDGNVY 143	
DB	79 GHNARFRDNL-----IQKRFPLG-----IWASYFKKPIGVLGIGAGPN---- 118	
QY	144 TAANDLSLQISRIHYGKFAEYKFRD-----AP--QDYEPILIRAKDKEGLM 190	
DB	119 ---NDSLUNYGIKRIINHAQFIT--VRDRESFDSLKHLSPSPVHETFDLIISKLREKT 174	
QY	191 KLITFTSVEETVRKV-----EKKAV-VFGQEVNLSDDNDN 226	
DB	175 EQLCOLKREAKDKIILVYHNHKKALEKFAESISLFLNPN 217	
RESULT 15		
Q8PUQ5	Q8PUQ5	PRELIMINARY; PRT; 1259 AA.
ID	Q8PUQ5;	
AC	Q8PUQ5;	
DT	01-OCT-2002 (TrEMBLrel. 22, Created)	
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	Hypothetical protein MW2277 (EC 2.7.3.-).	
GN	MW2277.	
OS	Methanosarcina mazei (Methanosarcina frisia).	
OC	Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;	
OC	Methanosarcinales; Methanosarcinaceae; Methanosarcina.	
OX	NCBI_TaxID=2209;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;	
RX	MEDLINE=22120827; PubMed=12125824;	
RA	Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,	
RA	Martinez-Arias R., Henne A., Wierzer A., Baeumer S., Jacobi C.,	
RA	Brueggemann H., Lienard T., Christmann A., Boencke M., Steckel S.,	
RA	Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,	
RA	Fritz H.-J., Gottschalk G.;	
RT	"The genome of Methanosarcina mazei: evidence for lateral gene transfer between Bacteria and Archaea.";	
RL	J. Mol. Microbiol. Biotechnol. 4:453-461(2002).	
DR	EMBL; AE013469; AAM31973.1; -.	
DR	GO; GO:0005524; F:ATP binding; IEA.	
DR	GO; GO:0016301; F:kinase activity; IEA.	
DR	GO; GO:0016740; F:transferase activity; IEA.	
DR	GO; GO:0000155; F:two-component sensor molecule activity; IEA.	
DR	GO; GO:0000160; P:two-component signal transduction system (p. .; IEA.	
DR	InterPro; IPR003594; ATPbind ATPase.	
DR	InterPro; IPR005467; His kinase.	
DR	InterPro; IPR001610; PAC_	
DR	InterPro; IPR000700; PAS-assoc.C.	
DR	InterPro; IPR000014; PAS_domain.	
DR	Pfam; PF02518; HATPase_C; 1.	
DR	Pfam; PF00785; PAC; 5.	
DR	Pfam; PF00989; PAS; 1.	
DR	TIGRFAMS; TIGR00229; sensory box; 6.	
DR	PROSITE; PS50109; HIS_KIN; 1.	
DR	PROSITE; PS50113; PAC; 5.	
DR	PROSITE; PS50112; PAS; 3.	
KW	Kinase; Transferase; Complete proteome; Hypothetical protein.	
SQ	SEQUENCE 1259 AA; 145131 MW; 9F523C6BF7A78440 CRC64;	

```

QY 9 PDSGNVTLASVREDLVRQEDTIIYGLI-----ERAKFPSNSHTYDEKYAQI--QGFCGS 61
Db 298 PDD-REVLINSIKKDPEQIRNIDYRIVLPDGEERTVTRVEITPDEENQPVRAEGIVQD 356
QY 62 LVEFVKNTE-ATQAKAGRYKNPEENAFPEENLPSPISVPSYFKOFLHPGAASININKSI 120
Db 357 ITE--LKRIEKALQEREQVR-----AFF-ENSMDAVL-----FASPDGTIHAANKAA 401
QY 121 WK--MYFKELLPLLATSGDDGNYAQTAANDLSLQSISSRIHYGKFAEV--KERDA-- 173
Db 402 CKTFGTEKEII-----RAGNGIVQD---SDPLKPSLEERNSTGRFKGEINHKKRQGTI 454
QY 174 -PDYBPLIRAKDKEGLMKLLTF-----TSVEETVRKRVEKXAVVF 213
Db 455 FPGEVSTAL-FKDKNGLMKIVIIIRDITERKKAEEVLKSEEHYRMLF 501

```

Search completed: September 25, 2004, 02:54:17  
Job time : 72 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 22:50:31 / Search time 4165 seconds  
(without alignments)  
10614.621 Million cell updates/sec

Title: US-10-624-061-15  
Perfect score: 1020  
Sequence: 1 gcacgagaggttaatttcaa.....atcttcaaaaaaaaaa 1020

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBank

- 1: gb\_ba:\*
- 2: gb\_hg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pt:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1020	100.0	1020	6	AR404688	AR404688 Sequence
2	465.6	45.6	525	6	AR404683	AR404683 Sequence
3	327.2	32.1	829	8	AY133840	AY133840 Arabidops
4	327.2	32.1	1006	6	AR236634	AR236634 Sequence
5	327.2	32.1	1006	6	AR236640	AR236640 Sequence
6	327.2	32.1	1006	6	AX343933	AX343933 Sequence
7	327.2	32.1	1006	6	AX464575	AX464575 Sequence
8	327.2	32.1	1006	6	AX464581	AX464581 Sequence
9	327.2	32.1	1006	8	L47355	L47355 Arabidopsis
10	327.2	32.1	1055	8	AY065238	AY065238 Arabidops
11	326.2	32.0	798	6	AX506277	AX506277 Sequence
12	326.2	32.0	993	6	AX343935	AX343935 Sequence
13	297.8	29.2	1000	8	L47356	L47356 Lycopersico
14	257.4	25.2	1231	6	AR404687	AR404687 Sequence
15	257.4	25.2	1250	8	AK068983	AK068983 Oryza sat
16	254.8	25.0	1274	8	AY089156	AY089156 Arabidops
17	254.8	25.0	1207	6	AR236633	AR236633 Sequence
18	254.8	25.0	1207	6	AR236639	AR236639 Sequence
19	254.8	25.0	1207	6	AX464574	AX464574 Sequence
20	254.8	25.0	1207	6	AX464580	AX464580 Sequence
21	254.8	25.0	1207	8	ATCHMUT	Z26519 A.thaliana
22	251	24.6	1005	6	AX506042	AX506042 Sequence
23	239.4	23.5	1223	6	AR404686	AR404686 Sequence
24	232.4	22.8	1435	8	AK101220	AK101220 Oryza sat
25	229.4	22.5	780	8	AR404689	AR404689 Sequence
26	229.4	22.5	780	8	BT009499	BT009499 Triticum
27	227.8	22.3	1043	8	AK069725	AK069725 Oryza sat
28	226	22.2	1348	8	AK105512	AK105512 Oryza sat
29	213	20.9	1143	8	AK117860	AK117860 Arabidops
30	212.6	20.8	951	8	BT005306	BT005306 Arabidops
31	211.4	20.7	1217	6	AR236635	AR236635 Sequence
32	211.4	20.7	1217	6	AR236641	AR236641 Sequence
33	211.4	20.7	1217	6	AX464576	AX464576 Sequence
34	211.4	20.7	1217	6	AX464582	AX464582 Sequence
35	211.4	20.7	1217	6	AF131219	AF131219 Arabidops
36	206.2	20.2	579	6	AR404684	AR404684 Sequence
37	173.4	17.0	8162	2	AC128660	AC128660 Medicago
38	132.4	13.0	246	8	AF012867	AF012867 Petroseli
39	129.8	12.7	394	8	PAJ004916	PAJ004916 Prunus av
40	123.2	12.1	2059	8	YSCAR07A	M24517 Saccharomyc
41	123.2	12.1	43776	8	SC9499X	Z49219 S.cerevisia
42	123.2	12.1	165536	8	SCCHRXVI	Z71255 S.cerevisia
43	99.8	9.8	843	6	BD266599	BD266599 Nucleic a
44	99.8	9.8	843	6	AX043758	AX043758 Sequence
45	99.8	9.8	1652	8	AF204738	AF204738 Fichla an

ALIGNMENTS

RESULT 1	AR404688	1020 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	Sequence 15 from patent US 6627798.				
DEFINITION	AR404688				
ACCESSION	AR404688				
VERSION	AR404688.1	GI:40153356			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 1020)				
AUTHORS	Falco, S.C., Famodu, O.O. and Lee, J.-M.				
TITLE	Aromatic amino acid biosynthetic enzymes				
JOURNAL	Patent: US 6627798-A 15 30-SEP-2003;				
FEATURES	Location/Qualifiers				





KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
COMMENT  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
COMMENT  
FEATURES  
source  
gene  
CDS  
3'UTR

FLI CDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 829)  
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W.,  
Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,  
Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R.,  
Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M.,  
Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P.,  
Southwick, A., Shinozaki, K., Sakurai, T., Satou, M., Seki, M., Shinn, P.,  
Arabidopsis Open Reading Frame (ORF) Clones  
Unpublished  
2 (bases 1 to 829)  
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W.,  
Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K.,  
Wong, C., Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Chen, H.,  
Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A.,  
Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M.,  
Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M.,  
Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and  
Theologis, A.  
Direct Submission  
Submitted (17-JUL-2002) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
The RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN  
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.  
The Salk, Stanford, PCEC (SSP) Consortium members constructed and  
sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K.,  
Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M.,  
Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,  
Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Chen, H.,  
Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Miranda, M.,  
Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M.,  
Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and  
Theologis, A.  
Yamada, K. (SSP/PCEC) and Seki, M. (RIKEN GSC) contributed equally  
to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP  
/PCEC) contributed equally to this work as PIs.  
Annotation is based on the January 2002 version of the Arabidopsis  
genome submitted to GenBank.  
Location/Qualifiers  
1..829  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="U18739"  
/ecotype="Columbia"  
/note="This clone is in pUNI 51."  
1..829  
/gene="At5g10870"  
1..798  
/gene="At5g10870"  
/codon\_start=1  
/evidence=experimental  
/product="putative chorismate mutase CM2"  
/protein\_id="PAM91774.1"  
/db\_xref="GI:22136860"  
/translation="MARVPESDSGSGCNVLSLDLIRSLRQEDTIVFLIERAKPP  
LNSPAPESRLDGSFSLTEFFVRETEIIQAKGVYEPPEENPFELEIPIHVSVP  
HYKGFALPKALSNVINKQIWDIYFKELLPLFVPGDDGNYPTAAADLACIALSER  
IHYKGFVAEVRPAQDPEAFIRAQDRALMLITFEKVEEMVKKVQKKAETFGQE  
VKFNSGYGDESKKRYKVDPLASRIYGEWLPIPLTKLVEVEYLLERLD"  
799..829

ORIGIN  
Query Match  
Best Local Similarity 32.1%; Score 327.2; DB 8; Length 829;  
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
/gene="At5g10870"  
QY 22 AATGCCCAACAGCAGACAAAGCTCTGATCTCTGGGAATGTGACACGCTAGCTTCTGT 81  
Db 6 AAGAGCTCTCGAATCGGATTCGGGTCTGGTGTCTCCAAATGACTGATCTTGTAAAT 65  
QY 82 GAGAGAGGATTTGGTTAGCAGAGAGATACCATCATTTATGGTCTCATGTAGAGAGCAA 141  
Db 66 CAGAGAACTGTTGATTAGCAGAGAGACACCATCTCTTCACTGATGATCGAGAGAGTAA 125  
QY 142 GTTCCTAGCAATCTCACACCTATGATGAAGATGCTCAAAATCCAGGTTTTTGG 201  
Db 126 GTTCCATCAATCTCTGCTTTCGAGGAATCTCGTGTCTAGATTTCTGAAGTTTCTC 185  
QY 202 CTCATTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAAGCTAAGGCTGGAAGATA 261  
Db 186 TTCTCTCACTGAGTTTTCGTCAGAGAGACAGAAATCATCAAGCTAAGTAGGAAGATA 245  
QY 262 CAAAACCTCTGAAGAAAGCCCTTCTCCAGAAATTTACCAACATCAATTTGTCCTC 321  
Db 246 TGAATACCCGGAAGAGAAATCTTCTCTCTGAGAACATCTCTCACTCGGTTTTTCTTAC 305  
QY 322 TTACTCTCTCAACAGTTTTTTCATCTCTGCTGCTTCAATTAACATAAACAAGTCCAT 381  
Db 306 GCACAATATCCTCGGCTTTGACACCTTAAGCTCTATCTGTATACATTAACAACAAT 365  
QY 382 CTGGAATATGATTTTCAAGAGATTAATCTCCATGCTTGTCTTCTGCTGATGATGGAA 441  
Db 366 CTGGGATATTTACTTTAAAGAAATTCCTTCTTGTGTTTGTCAACCTGGCGATGATGGCAA 425  
QY 442 CTATCGCAACCTGAGCTAATGACCTTTTCAATTTGAGTCCATCTCTAGAGGATTCATCA 501  
Db 426 CTATCATCACTGCTGCTGATGATCTGCTGCTTCAAGCTCTTTCAGAGAGATTCATCA 485  
QY 502 CTATGGAAGTTTGTAGCTGAGGTGAATTCAGGATGCTCTCTCAAGACTACGACCTTT 561  
Db 486 CTACGGTAAATTTGTAGCTGAGGTCAAAATTCAGAGATGCTCCACAAGATTACGACCTGC 545  
QY 562 AATTCGAGTATAGGATAAAGAGGATTTGATGAATTTGTGACATTTACAGGCTTTGAAGA 621  
Db 546 GATTCGCTCAGGATAGAGGCTTTGATGAAGCTGTTGAGCTTTGAGAAAGTAAAGA 605  
QY 622 GACGTTGAGGAAGAGATTTGAAAGAGGCTGTGTTTGGGAGGAGTGAATCTTAA 681  
Db 606 AATGGTTAAGAGAGAGTCCAGAGAGAGCAAGACGTTTGGACAAGAGTAAATTCATCA 665  
QY 682 -----CACTGATGACATGACAAATGAACCCGTAATTTGATTCATCATGCTGCTTCTAG 735  
Db 666 CTCGCTGATGCGATGAGAGTAAAGAAAGATATAAAGTGGATTCATTTGCTGCTCTCG 725  
QY 736 CTTGTACAAAATTTGGTGTATACCTCTCAAGAGAGGTTCAAGTTGAGTCACTCTTGGG 795  
Db 726 CATCTACGGGAAATGCTTATCCCTCTCACTAAGCTCGTTGAGTTGAGTTGATTTCTTCTAG 785  
QY 796 CGCTCTAGACTGAA 809  
Db 786 TCGTCTCGATTGAA 799  
RESULT 4  
AR236634  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Sequence 3 from patent US 6465217.  
AR236634  
AR236634.1  
GI:27280720  
Unknown.  
Unknown.

Unclassified.  
RESULT 5  
REFERENCE 1 (bases 1 to 1006)  
AUTHORS AR236640/c  
LOCUS AR236640  
DEFINITION Sequence 9 from patent US 6465217.  
TITLE AR236640  
ACCESSION AR236640  
VERSION AR236640.1 GI:27280726  
KEYWORDS  
SOURCE Unknown.  
FEATURES  
Location/Qualifiers  
1..1006  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 32.1%; Score 327.2; DB 6; Length 1006;  
Best Local Similarity 64.2%; Pred. No. 4e-66;  
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
QY 22 AATGGCCAAAGCAGCAGCAAAAGTCTGATTCCTGGAAATGTGTACACGCTAGCTTCTGT 81  
Db 69 AAGAGTCTTCGAATCGGATTCGGGTTCTGGTTCTTCCAAATGACTGAGTCTTGACTTAAT 128  
QY 82 GAGAGAGGATTTGTTAGGCAAGAGATACCATTTATGTTCTCATTTGAGAGAGCCAA 141  
Db 129 CAGAGAAATCGTTGATTAGGCAAGAGACACCATTCGTTTCAGTTGATCGAGAGCTAA 188  
QY 142 GTTCCCTPAGCAATTCACACCTATGATGAAAGATGCTCAATCCAGGGTTTTTGTGG 201  
Db 189 GTTCCCACTCAATTCCTGCTTTCGAGGAATCTGTTGCTAGATTCTGGAAGTTTCTC 248  
QY 202 CTCAATGGTGAATTTGTTTAAAGATACAGAGGCAATTCAGCTAAGCTTGGAGATA 261  
Db 249 TTCTCTCACTGAGTTTTCGTAGAGAGACAGAAATCATCCAGCTTAAGGTAGGAGATA 308  
QY 262 CAAAAACCCCTGAAGAAAAACCCCTTCTCCAGAAAAATTTACACCAATCAATTTGTGCCATC 321  
Db 309 TGAATACCCGGAAGAGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 368  
QY 322 TTACTCTTCAACAGTTTTCGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381  
Db 369 GCACAAATATCCATCGCTTTCGACCTTAAGGCTTCTATCTGTTAAATTAACAAACAAAT 428  
QY 382 CTGGAATAATGTTTCAAGAGTTACTTCCATTCGTTGCTTCTTCTTCTTCTTCTTCTTCTTCT 441  
Db 429 CTGGGATATTTACTTTAAGAAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 488  
QY 442 CTATGCGCAAACTGCAGCTAATGACCTTTCATTTATGTCAGTCACTCTCTAGAGGATTC 501  
Db 489 CTATCCATCAACTGCTGCTAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 548  
QY 502 CTATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCAAGACTACGAGCCTTT 561  
Db 549 CTACGGTAAATTTGTAGCTGAGTCAATTCAGAGATGCTCCAGAGATTACGAGCCTG 608  
QY 562 AATTCGAGCTAAGGATAAAGAGGATTGATGAAATTTGATCAATTTACAGCGTTTGAAGA 621  
Db 609 GATTTCGGCTCAGGATAGAGGCTTTGATGAAGCTTTGAGCTTTGAGAAAGTAGAAGA 668  
QY 622 GACGGTGAAGAGAGATTGAAAGAGGCTGTGGTTTGGCGGAGGAGTGAATCTTAA 681  
Db 669 AATGGTTAAGAGAGAGTGCAGAGAAAGCAGAAACGTTTTTGGACAGAGATTAATTTCAA 728  
QY 682 -----CAGTGATGACAAATGACAAATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735  
Db 729 CTCTGGCTATGGGATGAGAGTAAAGAAAGATTAAGTGGATCCATTTGCTTGGCTCTCG 788  
QY 736 CTTGTACAAAATTTGGGTATACCTCTCACAGAGGTTTCAAGTTGAGTACCTCTTGG 795  
Db 789 CATCTACGGGGATGCTTATCCCTCTCACTAAGCTCGTTGAGGTGATATCTTCTACG 848  
QY 796 CCGCTCTAGACTGAA 809  
Db 849 TCGTCTCGATTGAA 862

Unclassified.  
RESULT 5  
REFERENCE 1 (bases 1 to 1006)  
AUTHORS AR236640/c  
LOCUS AR236640  
DEFINITION Sequence 9 from patent US 6465217.  
TITLE AR236640  
ACCESSION AR236640  
VERSION AR236640.1 GI:27280726  
KEYWORDS  
SOURCE Unknown.  
FEATURES  
Location/Qualifiers  
1..1006  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 32.1%; Score 327.2; DB 6; Length 1006;  
Best Local Similarity 64.2%; Pred. No. 4e-66;  
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
QY 22 AATGGCCAAAGCAGCAGCAAAAGTCTGATTCCTGGAAATGTGTACACGCTAGCTTCTGT 81  
Db 938 AAGAGTCTTCGAATCGGATTCGGGTTCTGGTTCTTCCAAATGACTGAGTCTTGACTTAAT 879  
QY 82 GAGAGAGGATTTGTTAGGCAAGAGATACCATTTATGTTCTCATTTGAGAGAGCCAA 141  
Db 878 CAGAGAAATCGTTGATTAGGCAAGAGACACCATTCGTTTCAGCTTGCATCGAGAGAGCTAA 819  
QY 142 GTTCCCTPAGCAATTCACACCTATGATGAAAGATGCTCAATCCAGGGTTTTTGTGG 201  
Db 818 GTTTCCTCACTCAATTCCTGCTTTCGAGGAATCTCGTTGTCTAGATTCTTGAAGTTTCTC 759  
QY 202 CTCAATGGTGAATTTGTTTAAAGAAATACAGAGGCAATTCAGAGCTAAGGCTTGAAGATA 261  
Db 758 TTCTCTCACTGAGTTTTCGTAGAGAGACAGAAATCATCCAGCTTAAGGTAGGAGATA 699  
QY 262 CAAAAACCCCTGAAGAAAAACCCCTTCTCCAGAAAAATTTACACCAATCAATTTGTGCCATC 321  
Db 698 TGAATACCCGGAAGAGATCTTCTTCTTCTTGTGAGAAATCTCTCTCTCGGTTTTTCTTAC 639  
QY 322 TTACTCTTCAACAGTTTTCGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381  
Db 638 GCACAAATATCCATCGCTTTCGACCTTAAGGCTTCTATCTGTTAAATTAACAAACAAAT 579  
QY 382 CTGGAATAATGTTTCAAGAGTTACTTCCATTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441  
Db 578 CTGGGATATTTACTTTAAGAAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 519  
QY 442 CTATGCGCAAACTGCAGCTAATGACCTTTCATTTATGTCAGTCCATCTCTAGAGGATTC 501  
Db 518 CTATCCATCACTGCTGCTAGTATCTCGCTGTTTACAGCTCTTTCGAGAGGATTC 459  
QY 502 CTATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCAAGACTACAGGCTTT 561  
Db 458 CTACGGTAAATTTGTAGCTGAGTCAATTCAGAGATGCTCCACAGATTTACGAGCTG 399  
QY 562 AATTCGAGCTAAGGATAAAGAGGATTGATGAAATTTGATCAATTTTACAGCGTTTGAAGA 621  
Db 398 GATTTCGGCTCAGGATAGAGGCTTTTGTAGAGCTTTTGAAGTTTGAAGAGTAGAAGA 339  
QY 622 GACGGTGAAGAGAGATTGAAAGAGGCTGTGGTTTGGCGGAGGAGTGAATCTTAA 681  
Db 338 AATGGTTAAGAGAGAGTGCAGAGAAAGCAGAAACGTTTTTGGACAGAGATTAATTTCAA 279  
QY 682 -----CAGTGATGACAAATGACAAATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735

Db 278 CTCTGGCTATGCGATGAGATAAGAAAGATATAAAGTGGATCCATCTGCTGCTCG 219

Qy 736 CTGTGTACAAAATTTGGGTATACCTCTCACCAAGAGGTTTCAGGTTGAGTACCTCTTGGC 795

Db 218 CATCTACGGGAATGGCTTATCCCTCTCACTAAAGCTGTTGAGGTTGAGTATCTTCTACG 159

Qy 796 CCGTCTAGACTGAA 809

Db 158 TCGTCTCGAATTGAA 145

RESULT 6  
AX343933  
LOCUS AX343933 1006 bp DNA linear PAT 01-FEB-2002  
DEFINITION Sequence 3 from Patent WO0200901.  
ACCESSION AX343933  
VERSION AX343933.1 GI:18491972  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
AUTHORS Badur, R., Geiger, M., Kunze, I. and Sommer, S.  
TITLE Changing the fine chemical content in organisms by genetically  
modifying the snikimate pathway  
JOURNAL Patent: WO 0200901-A 3 03-JAN-2002;  
Sungene GmbH & Co. KGaA (DE)  
FEATURES  
source  
1. .1006  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
64. .861  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD22398.1"  
/db\_xref="GI:18491973"  
/db\_xref="REMBEL:CAD22398"  
/translation="MARVFSDSGSGSNVLSLDLRESLIROEDTIVFSLIRAKFP  
KNSFAFESCLDSGSFSSITEFFVETELIIQAKRYEYEPENPFLENPHSVPT  
HYPFALHFKALSNINKIWDIYFKELLPFVKPDGNYFPSTFASDLAQLARR  
IHVGKFAVEKFRADPOQYPAIRAODREALMKLLTPKVEEMVKRVOKKAEFTQGE  
VKFNSGVYDESKKVKYKVDPLLASRIYEWLIPLTKEVEVYLLRLD"

ORIGIN

Query Match 32.1%; Score 327.2; DB 6; Length 1006;  
Best Local Similarity 64.2%; Pred. No. 4e-66;  
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

Qy 22 AATGCCAAAGCAGAGACAAAGTCTCTGATTCTGGGAATGTGTACACGCTAGTCTCTGT 81

Db 69 AAGAGTCTCGAATCGGATTCGGGTTCTGGTTGTCCAAATGTACTGATCTTGACTTAAT 128

Qy 82 GAGAGAGATTGGTTAGGCAAGAGATACCATATTATGTTCTCAATTGAGAGAGCCAA 141

Db 129 CAGAGATCGTTGATTAGGCAAGAGACACCATCGTCTCAGCTTGATCGAGAGAGTAA 188

Qy 142 GTTCCCTAGCAATCTCACCTATGATGAAGATGCTCAATCCAGGTTTCTGG 201

Db 189 GTTCCCACTCAATCTCTCGTTTCGAGGAATCTGTTGTCTAGATTCTGGAAGTTCTC 248

Qy 202 CTCATTGGTGAATTTGTTTAAAGATACAGAGGCCATTCAAGCTAAGCGTGAAGATA 261

Db 249 TTCTCTCACTGATTTTTCGTCAGAGAGACAGAAATCATCCAACTAAGCTAGGAGATA 308

Qy 262 CAAMAACCTGAAGAAAGCGCTTCTCCAGAAAATTTACCACCATCAATTGCGCATC 321

Db 309 TGAATACCCGAGAGAAATCTTTCTCTTCTGAGAAATCTCCCTCAGCTGTTTCTCTAC 368

Qy 322 TTACTCTTCAACAGATTTTTCATCTCTGCTGCTTCAATTAACATAACAAAGTCCAT 381

Db 369 GCACAAATATCATCGGCTTTTCACCCCTAAGGCTCTATCTGTTAACATTTAACAACAAT 428

Qy 382 CTGGAATAATGTTTCAAAAGAGTTACTTCCATTGCTTGTCTACTTCGGGTGATGATGGCAA 441

Db 429 CTGGGATATTTACTTTAAAGAAATGCTTCCCTTTGTTGTTCAAACTGGCGATGATGGCAA 488

Qy 442 CTATGCGCAAACCTGAGCTAATGAGCTTTTCATTTGCAAGTCCATCTCTAGAGGATTCOA 501

Db 489 CTATCCATCAACTGCTGCTAGTATGATCGCTGTTTACAAGCTCTTTCCGAGAGGATTCOA 548

Qy 502 CTATGGAAGATTTGTAGCTGAGTGAGTGAATTTCAAGGATGCTCTCAAGACTACGAGCTTT 561

Db 549 CTACGGTAATTTGTAGCTGAGTCAATTCAGAGATGCTCCACAAGATTACGAGCTGTC 608

Qy 562 AATTGAGCTAAGGATAAAGAGATTGATGAAATTTGTTGACATTTACAAGCGTTGAAGA 621

Db 609 GATTCGCGCTCAGGATAGAGAGCTTTGATGAAGCTGTTGACGTTTGAGAAAGTAGAAGA 668

Qy 622 GACGCTGAGGAAGAGATTGAAAGAGGCTGTGCTGTTTGGCAGGAGTCAATCTTAA 681

Db 669 AATGGTTAAGAGAGAGTGCAGAAAGAGACAAAGCTTTGGACAAAGATGAAATTTCAA 728

Qy 682 -----CACTGATGACAATGACAATCAAAACCGTAAATTTGATCCATCAGTGGCTTTAG 735

Db 729 CTCTGCTATGCGATGAGAGTAAAGAGATATAAAGTGGATCCATTGCTTGCCTCTCG 788

Qy 736 CTGTACAAAATTTGGTGATACCTCTCACCAAGAGGTTTCAAGTTGAGTACCTCTTGGC 795

Db 789 CATCTACGGGAATGCTTATCCCTCTCACTAAGCTCGTTGAGTTGAGTATCTTCTTACG 848

Qy 796 CCGTCTAGACTGAA 809

Db 849 TCGTCTCGAATTGAA 862

RESULT 7  
AX464575  
LOCUS AX464575 1006 bp DNA linear PAT 16-JUL-2002  
DEFINITION Sequence 3 from Patent WO0202798.  
ACCESSION AX464575  
VERSION AX464575.1 GI:21899370  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
AUTHORS Boyes, D.C., Davis, K.R., Woessner, J.P., Gorlach, J., Hamilton, C.M.,  
Hoffman, N.E., Kloti, A.S., Zayed, A., Ascenzi, R.A., Allen, K.,  
Mulpuri, R. and Kjemtrup, S.  
TITLE Methods and compositions for the modulation of chorismate synthase  
and chorismate mutase expression or activity in plants  
JOURNAL Patent: WO 0202798-A 3 10-JAN-2002;  
Paradigm Genetics Inc. (US)  
FEATURES  
source  
1. .1006  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN

Query Match 32.1%; Score 327.2; DB 6; Length 1006;  
Best Local Similarity 64.2%; Pred. No. 4e-66;  
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

Qy 22 AATGCCAAAGCAGAGACAAAGTCTCTGATTCTGGGAATGTGTACACGCTAGTCTCTGT 81

Db 69 AAGAGTCTCGAATCGGATTCGGGTTCTGGTTGTCCAAATGTACTGATCTTGACTTAAT 128

Qy 82 GAGAGAGATTGGTTAGGCAAGAGATACCATCAATTTATGTTCTCAATTGAGAGAGCCAA 141

Db 129 CAGAGATCGTTGATTAGGCAAGAGACACCATCGTCTCAGCTTGATCGAGAGAGTAA 188

Qy 142 GTTCCCTAGCAATCTCACCTATGATGAAGATGCTCAATCCAGGTTTCTGG 201

Db 189 GTTCCCACTCAATCTCTCGTTTCGAGGAATCTGTTGTCTAGATTCTGGAAGTTCTC 248

Qy 202 CTCATTGGTGAATTTGTTTAAAGATACAGAGGCCATTCAAGCTAAGCGTGAAGATA 261

Db 249 TTCTCTCACTGATTTTTCGTCAGAGAGACAGAAATCATCCAACTAAGCTAGGAGATA 308

Qy 262 CAAMAACCTGAAGAAAGCGCTTCTCCAGAAAATTTACCACCATCAATTGCGCATC 321

Db 309 TGAATACCCGAGAGAAATCTTTCTCTTCTGAGAAATCTCCCTCAGCTGTTTCTCTAC 368

Qy 322 TTACTCTTCAACAGATTTTTCATCTCTGCTGCTTCAATTAACATAACAAAGTCCAT 381

QY 142 GTTCCCTAGCAATTCACACCTATGATGAAAGTAGTGTCAAAATCCAGGGTTTGTGG 201  
 |||||  
 Db 189 GTTCCCACTCAATTCCTCTCTTTCAGGGAATCTGTTGTCTAGATTCTGGAAGTTTCTC 248  
 |||||  
 QY 202 CTCAATGGTGAATTTGTTTGAAGATACAGAGGCCTTCAAGCTAAGCTGGAGATA 261  
 |||||  
 Db 249 TTCTCTCACTGAGTTTTTTCGTACAGAGACAGAAATCATCCAGCTAAGGTAGGAGATA 308  
 |||||  
 QY 262 CAAAAACCCCTGAAGAAACCCCTTCTTCCAGAAAAATTTACCAACCATCAATTTGTGCAATC 321  
 |||||  
 Db 309 TGAATACCGGAGAGATCTTCTTCTTGAAGACATTCCTCACTCGTGTTCCTTAC 368  
 |||||  
 QY 322 TTACTCTTCAACAGTTTTTGATCTCTGTGTCTGCTTCAATTTAAACATAACAGTCCAT 381  
 |||||  
 Db 369 GCACAAATATCCATCGGCTTTGACCCCTAAGGCTCTATCTGTTAACATTTAAACAAAT 428  
 |||||  
 QY 382 CTGMAAATGATTTCAAGAGTTACTTCCATTTGCTTGTACTTTCGGGTGATGGCAA 441  
 |||||  
 Db 429 CTGGATTTTACTTTAAAGATTGCTTCTTGTGTTCAACCTGGCGATGGCAA 488  
 |||||  
 QY 442 CTATGCAAACTGACGCTAATGACCTTTCATTTATTCAGTCAATCTCTAGAGATTCA 501  
 |||||  
 Db 489 CTATCCATCACTGCTGCTAGTATCTCGCTGTTTACAAAGTCTTTCGAGAGATTCA 548  
 |||||  
 QY 502 CTATGGAAGTTGTAGCTGAGGTGAATTCAGGATGCTCTCAAGACTACGAGCCTTT 561  
 |||||  
 Db 549 CTACGGTAAATTTGTAGTGAAGTCAATTCAGAGATGCTTCCACAGATTACGAGCTGC 608  
 |||||  
 QY 562 AATTGCGAGCTAAGGATAAGAGAGATTGATGAATTTGACATTTTACAAGCGTTGAAGA 621  
 |||||  
 Db 609 GATTGCGCTCAGGATAGAGAGCTTTTATGAAGCTGTTGACGTTTGAAGAAAGTAGAGA 668  
 |||||  
 QY 622 GAGGTGAGGAAGAGATTGAAGAGAGCTGTGTTTGGCAGGAGTCAATCTTAA 681  
 |||||  
 Db 669 AATGGTTAAGAGAGAGTGCAGAGAGAAAGCAGAAACGTTTGGACAGAAAGTAAATCAA 728  
 |||||  
 QY 682 -----CAGTGTAGCAATGACATGAACCCGTAATTTGATCCATCAGTGGCTTCTAG 735  
 |||||  
 Db 729 CTCTGCTATGCGGATCAGAGTAGAAGATATAAGTGGATCCATTGCTTGCCTCTCG 788  
 |||||  
 QY 736 CTCTGACAAAATTTGGTGTATCTCTCACCAGAGAGTTGAGTTGATCTCTTGGG 795  
 |||||  
 Db 789 CATCTACGGGGAATGCTTATCCCTCTCACTAGCTCGTTGAGGTTGAGTATCTTCTAG 848  
 |||||  
 QY 796 CCGTCTAGACTGAA 809  
 |||||  
 Db 849 TCGTCTCGATTGAA 862  
 |||||

RESULT 8  
 AX464581/c 1006 bp DNA linear PAT 16-JUL-2002  
 LOCUS Sequence 9 from Patent WO0202798.  
 DEFINITION AX464581

AX464581.1 GI:21899376

ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
 AUTHORS Boyes, D.C., Davis, K.R., Woessner, J.P., Gorlach, J., Hamilton, C.M.,  
 Hoffman, N.E., Kloti, A.S., Zayed, A., Ascenzi, R.A., Allen, K.,  
 Mulpuri, R. and Kjettrup, S.

TITLE Methods and compositions for the modulation of chorismate synthase  
 and chorismate mutase expression or activity in plants

JOURNAL Patent: WO 0202798-A 9 10-JAN-2002;

Paradigm Genetics Inc. (US)

FEATURES Location/Qualifiers

source 1..1006 /organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"  
 /db\_xref="taxon:3702"

# ORIGIN

Query Match 32.1%; Score 327.2; DB 6; Length 1006;  
 Best Local Similarity 64.2%; Pred. No. 4e-66;  
 Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
 QY 22 AATGCCAAAGCAGCAGACAAAGTCTCTGATTCTGGGAATGTGTACACGCTAGCTTCTGT 81  
 Db 938 AAGAGTCTTCGAATCGGATTCTGGTCTGCTTCTCAATGTACTAGTCTTGACATTAAT 879  
 |||||  
 QY 82 GAGAGAGATTTGGTTAGGCAAGAGGATACCATCATTTATGTCTCATTTGAGAGGCCAA 141  
 |||||  
 Db 878 CAGAGAAATCGTTGATTAGGCAAGAGACACCATCGTCTTCAGCTTGAATCAGAGAGCTAA 819  
 |||||  
 QY 142 GTTCCCTAGCAATTTCTCACACCTATGATGAAAGTAGTGTCTCAATCCAGGGTTTTTGTGG 201  
 |||||  
 Db 818 GTTTCCACTCAATTTCTCTGCTTTTCGAGGAATCTCGTTGTCTAGATTCTGGAAGTTTCTC 759  
 |||||  
 QY 202 CTCAATGGTGAATTTGTTTGAAGATACAGAGGCATTCAGCTAAGCTCGAGAGATA 261  
 |||||  
 Db 758 TTCTCTCACTGAGTTTTTTCGTACAGAGACAGAAATCATCCAGCTAAGGTAGGAGATA 699  
 |||||  
 QY 262 CAAAAACCCCTGAAGAAACCGCTTCTTCCAGAAAAATTTACCAACCATCAATTTGTGCAATC 321  
 |||||  
 Db 698 TGAATACCCGAGAGAGATCTTCTTCTTGAAGACATTCCTCACTCGTGTTCCTTAC 639  
 |||||  
 QY 322 TTACTCTTCAACAGTTTTTGTGATCTGTGTCTGCTTCAATTTAAACATAACAGTCCAT 381  
 |||||  
 Db 638 GCACAAATATCCATCGGCTTTGCACCTAAGGCTCTATCTGTTTAAACATTAACAAACAAAT 579  
 |||||  
 QY 382 CTGGAATAATGTTTCAAGAGTTTACTTCCATTTGCTTGTCTACTTCGGTGTATGATGCAA 441  
 |||||  
 Db 578 CTGGATATTACTTTAAAGAAATGCTTCTTCTTGTGTTGTTCAACCTGGCGATGATGCAA 519  
 |||||  
 QY 442 CTATGCGCAAACTGACGCTAATGACCTTTTCATTTATTCAGTCCATCTCTAGAAAGATTCA 501  
 |||||  
 Db 518 CTATCCATCACTGCTGCTAGTGAATCTCGCTCTTTTACAAGCTCTTTTCGAGAGAGATTCA 459  
 |||||  
 QY 502 CTATGGAAAGTTGTAGCTGAGGTGAATTTACGGATGCTCCTCAAGACTACGAGCCTTT 561  
 |||||  
 Db 458 CTACGGTAAATTTGTAGCTAGGTCAATTCAGAGATGCTCCACAAGATTACAGGCTGC 399  
 |||||  
 QY 562 AATTGCGAGCTAAGGATAAGAGAGATTGATGAATTTGTTGACATTTTCAAGCGTTGAAGA 621  
 |||||  
 Db 398 GATTTCGGCTCAGGATAGAGAGGCTTTGATGAAGCTGTTGACGTTTGAAGAAAGTAGAGA 339  
 |||||  
 QY 622 GACGTCAGGAAGAGAGTTGAAAAGAGAGCTGTGTTGTTGGCAGGAAGTGAATCTTAA 681  
 |||||  
 Db 338 AATGGTTAAGAGAGAGTGCAGAGAAAGCAGAAACGTTTGGACAGAAAGTAAATTTCAA 279  
 |||||  
 QY 682 -----CAGTGTAGCAATGACATGAACCCGTAATTTGATCCATCAGTGGCTTCTAG 735  
 |||||  
 Db 278 CTCTGCTATGCGGATCAGAGTAGAAGATATAAGTGGATCCATTGCTTGCCTCTCG 219  
 |||||  
 QY 736 CTCTGACAAAATTTGGTGTATCTCTCACCAGAGAGTTGAGTTGATCTCTTGGG 795  
 |||||  
 Db 218 CATCTACGGGGAATGCGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTTCTAG 159  
 |||||  
 QY 796 CCGTCTAGACTGAA 809  
 |||||  
 Db 158 TCGTCTCGATTGAA 145  
 |||||

# RESULT 9

LOCUS L47355

DEFINITION Arabidopsis thaliana

ACCSSION L47355

VERSION L47355.1

KEYWORDS chorismate mutase; mutase.

SOURCE Arabidopsis thaliana (thale cress)

L47355 1006 bp mRNA linear PLN 13-AUG-1999  
 Arabidopsis thaliana chorismate mutase mRNA, complete cds.

ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 1006)  
AUTHORS Eberhard, J., Enzler, T., Epple, P., Felix, G., Raesecke, H.R.,  
Arnheim, N. and Schmid, J.  
TITLE Cytosolic and plastidic chorismate mutase isozymes from Arabidopsis  
thaliana: molecular characterization and enzymatic properties  
JOURNAL Plant J. 10 (5), 815-821 (1996)  
MEDLINE 97111372  
PUBMED 8953244  
COMMENT GSDB:S:46595  
FEATURES  
Location/Qualifiers  
1..1006  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/cldes="catCM2"  
64..861  
/EC\_number="5.4.99.5"  
/note="putative"  
/codon\_start=1  
/product="chorismate mutase"  
/protein\_id="AAD48922.1"  
/db\_xref="GI:5732016"  
/translation="MARVPSDSGSGSNVSLDLRLSRLRQEDTIVSLIERAKFP  
LNSPAPESCLDSGSSLSLEFVRETEIIQAKVGRYEPENPFLENPHSVFPT  
HKYPSALHPKALSINNKIWDIYFKELLEFLFKPDGDNYPSTASDLACIQALSR  
ITKNGKVAEYKFDAPQDEPAIQAQDRALMKLLTFKVEEMVKVRVOKAETFOQE  
VFNKSGYGSDESKKKYKVDPLLASRIYGEWLIPLTKLVEVEYLLRLD"  
1006  
polya\_site  
/note="putative"  
ORIGIN  
Query Match 32.1%; Score 327.2; DB 8; Length 1006;  
Best Local Similarity 64.2%; Pred. No. 4e-66;  
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
Qy 22 AATGCCCAAGCAGCAGCAAAAGTCTCTGATTCGGGAAGTGTACACGCTAGCTCTGT 81  
Db 69 AAGAGTCTTCAATCGGAATCGGGTCTGGTGTTCATGATGCTAGTCTTCACTTAAT 128  
Qy 82 GAGAGAGGATTTGGTTAGGCAAGAGGATACCATTTATGGTCTCATTCAGAGAGCCAA 141  
Db 129 CAGAGAATCGTTGATTAGGCAAGAGACACCATCTTCAGCTTGATCGAGAGAGCTAA 188  
Qy 142 GTTCCCTAGCAATCTCACACCTATGATGAAGATGCTCAATCCAGGCTTTTGTGG 201  
Db 189 GTTTCACCTCAATCTCTCTGTTTCGAGGAATCTCGTGTCTAGATTCGGAAGTTCTC 248  
Qy 202 CTCATTGGTGAATTTGTTTGAAGATACAGAGGCCATTCAGCTAAGGCTGGAAGATA 261  
Db 249 TTCTCTCAGTGTTTTCGTGAGAGAGAGAGAAATCATCCAAGCTAAGGTAGAGATA 308  
Qy 262 CAAAAACCTGAAGAAACGCTTTCTTCCAGAAATTTACCAACCATCAATTTGCCATC 321  
Db 309 TGAATACCGGAGAGAAATCTTTCTTCTTCTTGAAGAAATCTCTCACTCGGTTTTCTAC 368  
Qy 322 TTACTCTTCAACAGAGTTTTCATCTCTGCTGCTTCAATTAACATAACAAAGTCCAT 381  
Db 369 GCACAAATATCCATCGGCTTTCACCTTAAGGCTCTATCTGTATCAATCAACAAAT 428  
Qy 382 CTGGAATATGTAATTCAGAGGTTACTTCCATTCCTTCTGCTTCTGGGTGATGAGCAA 441  
Db 429 CTGGATATTACTTTAAAGAAATGCTTCTCTTCTTGTCTCAACCTGGCATGAGCAA 488  
Qy 442 CTATGCGCAACTCAGCTAATGACCTTTCATTTATTCAGCTCCATCTCTAGAGGATTC 501  
Db 489 CTATCCATCACTCTGCTAGTATCTCGGCTGTTTCAAGCTCTTTTCGAGAGGATTC 548  
Qy 502 CTATGGAAGTTTGTAGTCTGAGTGAATTCAGGGATGCTCTCAAGACTACGAGCTTT 561

Db 549 CTACGGTAAATTTGTAGCTGAGGTCAAATTCAGAGATCTCCACAAGATTACGACCTGC 608  
Qy 562 AATTCAGCTAGGATAAAGAGGATTTGATGAATTTGTTGACATTTACAAGCGTTGAAGA 621  
Db 609 GATTCGCGCTCAGGATAGAGAGGCTTTGATGAGCTGTTGACGTTTGAGAAAGTAGAAGA 668  
Qy 622 GACGCTGAGGAGAGAGTTGAAAGAGAGCTGTGTTGTTGGGAGGAGTGAATCTTAA 681  
Db 669 AATGTTAAGAGAGAGTGCAGAGAGAAAGCAAGACGTTTGACAGAAAGTAAATCAA 728  
Qy 682 -----CAGTGATGACAATGACAATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735  
Db 729 CTCTGCTATCGGATGAGAGTAAAGAGATTAAGTGGATCCATTTGCTGCTCTCG 788  
Qy 736 CTTGTACAAAATTTGGGTGATACCTTCCACAGAGAGTTGAGTTGAGTCTCTTCCG 795  
Db 789 CATCTACGGGAATGCTTATCCCTCTCACTAAGCTCGTTGAGTTGATATCTTCTACG 848  
Qy 796 CCGTCTAGACTGAA 809  
Db 849 TCGTCTGATGAA 862

RESULT 10  
LOCUS AY065238  
DEFINITION Arabidopsis thaliana putative chorismate mutase CM2 (At5g10870)  
mRNA, complete cds.  
ACCESSION AY065238  
VERSION AY065238.1 GI:17529007  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 1055)  
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,  
Davis, R.W., Ecker, J.R. and Theologis, A.  
Arabidopsis Full Length cDNA Clones  
Unpublished  
2 (bases 1 to 1055)  
Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M.,  
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (03-DEC-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA

TITLE Arabidopsis Full Length cDNA Clones  
JOURNAL Unpublished  
AUTHORS Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M.,  
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (03-DEC-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA

TITLE Arabidopsis Full Length cDNA Clones  
JOURNAL Unpublished  
AUTHORS Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M.,  
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (03-DEC-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA

TITLE Arabidopsis Full Length cDNA Clones  
JOURNAL Unpublished  
AUTHORS Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M.,  
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (03-DEC-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA

TITLE Arabidopsis Full Length cDNA Clones  
JOURNAL Unpublished  
AUTHORS Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M.,  
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (03-DEC-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA

TITLE Arabidopsis Full Length cDNA Clones  
JOURNAL Unpublished  
AUTHORS Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M.,  
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (03-DEC-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA

Yamada, K. (SSP/PGE) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGE) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

## FEATURES

## source

1. 1055  
Location/Qualifiers

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="RAF109-41-E09 (R18739)"

/ecotype="Columbia"

/notes="This clone is in a modified pBluescript vector (FUC-1) as a BamHI/XhoI insert."

1. 1055

/gene="At5g10870"

1. 98

/gene="At5g10870"

## misc\_difference 2

/gene="At5g10870"

/notes="compared to genomic sequence"

/replaces="t"

99. 896

/gene="At5g10870"

/codon\_start=1

/evidence=experimental

/product="putative chorismate mutase CM2"

/protein\_id="AAL38714.1"

/db\_xref="GI:17529008"

/translation="MARVPSDSGSGSNVSLDLRLSRLPQEDTIVFSILIERAKFP

LSNPARESRCLDSGFSLSLTFVRETEIIQAKGRVHYEPENFFLENPHSVFPT

HKYPSALHPKALSNINKQIWDIYFKLELPLFKPGDDGNYFSTPAASDLACLQALSR

IHYGKFAVBEKDDYQDPEAPRAQDREALMKLJTFKEVEMVKRVOKAETFGQE

VFNISGYGDESKKXVDFLLASRIYGEWLIPLTKLIVEVEYLLRLD"

897. 1055

/gene="At5g10870"

## ORIGIN

Query Match 32.1%; Score 327.2; DB 8; Length 1055;

Best Local Similarity 64.2%; Pred. No. 4e-66;

Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 AATGGCCAAAGCAGCAGACAAAGTCCTGATTCTGGGAATGTGTACACGCTAGCTTCTGT 81

DB 104 AAGAGTCTCGAATCGAATCGGGTCTGGTGTCTCCAAATGACTGAGTCTTGACTTAAT 163

QY 82 GAGAGAGGATTTGGTTAGGCAAGAGGATACCATCAATTTATGGTCTCAATGAGAGGCCAA 141

DB 164 CAGAGAATCGTTGATTAGGCAAGAGACACCATCGTCTTCAGCTTGATCGAGAGAGCTAA 223

QY 142 GTTCCCTAGCAATCTCACACCTATGATGAAGATGCTCAATCCAGGTTTGTGG 201

DB 224 GTTCCACTCAATCTCCGCTTCAGGAATCTCGTGTCTAGATCTCGGAATTTCTC 283

QY 202 CTCATTTGGTGAATTTGTTTGAAGAAATACAGAGGCCATTCAGCTTAAGCTTGAAGATA 261

DB 284 TTCTCTCAGTGATTTTTCGTGAGAGACAGAAATCATCAAGCTTAAGTAGGAAGATA 343

QY 262 CAAAACCTTGAGAAAACGCTTCTTCCAGAAAATTTACCAACCATCAATTTGTCGCATC 321

DB 344 TGAATACCCGGAAGAGAAATCTTCTCTCTTGAGAACATTCCTCACTCGGTTTTTCCATC 403

QY 322 TTAATCTCTCAACAGTTTTCATCTGCTGCTTCAATTAACATAAACAAGTCCAT 381

DB 404 GCACAAATATCCATCGGCTTCACCCCTAAGGCTCTATCTGTTAATTAACAACAAT 463

QY 382 CTGGAAAATGATTTCAAGAGTTACTTCGATGCTTGCTTACTCTGGGTGATGATGGCAA 441

DB 464 CTGGGATATTTACTTTTAAAGAAATTTGCTCTCTTTGTTTGTCAAAACCTGGCGATGATGGCAA 523

QY 442 CTATGGCAAACCTGCAGCTAATGACCTTTTCATTTATTTGCAGTCCATCTCTAGAAGGATTTCA 501

DB 524 CTATCCATCAACTGCTGTAGTATCTCGGCTGTGTACAAAGCTCTTTTCGAAAGGATTTCA 583

QY 502 CTATGGAAAGTTTGTAGCTGAGGTGAATTCAGGATGCTCTCAAGACTTACGAGCTTTT 561

DB 584 CTACGGTAAATTTGTAGCTGAGGTCAAAATTCAGAGATGCTCCACAGATTACGAGCTGC 643

QY 562 AATTCGAGCTAAGGATAAAGAAGATTGATGAATTTGTTGACATTTTCAAGCGTTGAAGA 621

DB 644 GATTTCGGCTCAGGATAGAGAGGCTTTGATGAAGCTGTTCACGTTTTCAGAAAAGTAGAAGA 703

QY 622 GACGTTGAGGAGAGAGTTTGAAGAAGAGCTGTGGTGTGTTGGCAGGAGTGAATCTTAA 681

DB 704 AATGGTTAAAGAGAGAGTGCAGAGAAGCAGAAACGTTTGGACAGAGATAAATTTCAA 763

QY 682 -----CAGTGATGACAATGACAATGAAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735

DB 764 CTCTGGCTATGGCGATGAGAGTGAAGAAGTATAAAGTGATCCATTCCTTCCCTCTCG 823

QY 736 CTGTACAAAATTTGGTGATACCTCTCACCAGAGGTTTCAGTTGAGTACCTCTTGG 795

DB 824 CATCTACGGGAATGGCTTATCCCTCTCACTAAGCTGTTGAGTTGAGTATCTTCTACG 883

QY 796 CCGTCTAGACTGAA 809

DB 884 TCGTCTCGATTGAA 897

## RESULT 11

AX506277

LOCUS

AX506277

DEFINITION

AX506277

ACCESSION

AX506277.1

VERSION

AX506277.1

KEYWORDS

ORGANISM

Arabidopsis thaliana (thale cress)

REFERENCE

AUTHORS

TITLE

JOURNAL

THE SCRIPPS RESEARCH INSTITUTE (US); Syngenta Participations AG (CH)

FEATURES

source

i. 798

/organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"

/db\_xref="taxon:3702"

ORIGIN

Query Match 32.0%; Score 326.2; DB 6; Length 798;

Best Local Similarity 64.2%; Pred. No. 7e-66;

Matches 509; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 AATGGCCAAAGCAGCAGACAAAGTCCTGATTCTGGGAATGTGTACACGCTAGCTTCTGT 81

DB 6 AAGAGTCTTCGAATCGAATTCGGGTTCTGGTGTCTCCAAATGACTGAGTCTTGACTTAAT 65

QY 82 GAGAGAGGATTTGGTTAGGCAAGAGGATACCATCAATTTATGGTCTCATTTGAGAGGCCAA 141

DB 66 CAGAGAATCGTTGATTAGGCAAGAGACACCATCGTCTTCAGCTTGCATCGAGAGAGCTAA 125

QY 142 GTTCCCTAGCAATTTCTCACACCTATGATGAAGATGCTCAATCCAGGTTTGTGG 201

DB 126 GTTTCCTCAATCTCTCTGTTTCGAGGAATCTCGTGTCTAGATTCGGAGTTTCTC 185

QY 202 CTCATTTGGTGAATTTGTTTGAAGAAATACAGAGGCCATTCAGAGCTTGAAGATA 261

DB 186 TTCTCTCACTGAGTTTTCGTTCAGAGAGACAGAAATCATCCAAAGCTAAGGTAGGAAGATA 245

```
QY 262 CAAAACCCCTGAAGAAACGCTTCTTCCAGAAAAATTTACCAACATCAATTTGGCCATC 321
Db 246 TGAATACCCGGAAGAGATCTTCTTCTTCTTGAGAACATCTCTCACTCGGTTTTCCTAC 305
QY 322 TTACTCCTTCAAACAGTTTTTGCATCTCTGCTGCTCTCAATTAACATAACAAGTCCAT 381
Db 306 GCACAAATATCCATCGGCTTTGCACCCCTTAAGGCTCTATCTGTTAAACATTAACAAACAAT 365
QY 382 CTGGAATATGATTTCAAAGAGTTACTTCCATTCCTTCTGCTACTTCCGGGTGATGATGCAA 441
Db 366 CTGGGATATTTACTTTAAAGAAATTCCTTCTTCTTCTGTCAAACCTGGCGATGATGCAA 425
QY 442 CTATGCGCAACTCGAGCTAATGACCTTTCAATTTGAGAGTCCATCTCTAGAGGATCA 501
Db 426 CTATCCATCAACTGCTGCTAGTATCTGCGCTGTTTAAAGCTCTTTTCGAGAGGATCA 485
QY 502 CTATGGAAGTTTCTAGCTGAGGTGAAATTCAGGATGCTCTCAAGACTACGAGCCTTT 561
Db 486 CTACGGTAATTTGCTAGCTGAGTCAATTCAGAGATGCTCCCAAGATTAACAGCCTGC 545
QY 562 AATTCGAGCTAAGGATTAAGAGGATGATGAATTTGACATTTTACAGCGTTGAAGA 621
Db 546 GATTCGGGCTCAGGATAGAGGCTTTGATGAAGCTGTTGACGTTTGAGAAAGTAGAAGA 605
QY 622 GACGTTGAGGAGAGGTTGAAGAAGGCTGCTGTTTGGCAGGAGTGAATCTTAA 681
Db 606 AATGTTTAAGAGAGAGTGCAAGAAGAGCAAGACGTTTGGACAGAAGTAAATTCAA 665
QY 682 -----CAGTGATGACATGACATGAATGAACCGTAAATTTGATCCATCAGTGGCTTAG 735
Db 666 CTCTGGCTATGGCATGAGAGTAAGAAGATATAAGTGAATCAATTTGCTTGGCTCTCG 725
QY 736 CTCTGACAAATTTGGTGATACCTCTCACCAGGAGGTTGAGTTGAGTACCTCTTGG 795
Db 726 CATCTACGGGAATGGCTTATCCCTCACTAAGCTCGTTGAGTTGAGTATCTTACG 785
QY 796 CCGTCTAGACTGA 808
Db 786 TCGTCTCGATTGA 798

RESULT 12
AX343935
LOCUS
DEFINITION Sequence 5 from Patent WO0200901.
ACCESSION AX343935
VERSION AX343935.1 GI:18491974
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1 Badur,R., Geiger,M., Kunze,I. and Sommer,S.
AUTHORS Changing the fine chemical content in organisms by genetically
TITLE modifying the shikimate pathway
JOURNAL Patent: WO 0200901-A 5 03-JAN-2002;
Sungene GmbH & Co. KGaA (DE)
FEATURES
Location/Qualifiers
1..993
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Chimäre Nukleins.: Transilpt. der plastid.
Chorismatm.+kod. Sequenz der cytosolischen
Chorismatmutase"
1..993
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD22399.1"
/db_xref="GI:18491975"
/db_xref="RENTREMBL:CAD22399"

CDS
```

```
/translation="MRSSCCSAIGGFFDHRRELSTSTPSTLLPLPSTKSFVSRCV
LPQKPSRGTSSVHVMVTAMARVPSDSCGCSNLSLDLIRSLRQEDTIVFSL
IFRAKPLNPAFESRCLDSGSFSLTFFVRETEIIOAKVRYEVEPEENPFLENI
PHSVPTHKYPSALHPKALSUNIKQIWDIYFKELLPLVFKPDGDNPFSTAASDIAC
IQALSRHYGKFAVEVFRDPAQDIEPAIRQDREALKLITFKYEMWAKRVQKK
AETFGQVKFNISGYSKSKYKVDPLLASRIYGEWLLPLTKLVEVEYLLRLD"

ORIGIN
Query Match 32.0%; Score 326.2; DB 6; Length 993;
Best Local Similarity 64.2%; Pred. No. 6.8e-66;
Matches 509; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 AATGGCAAGCAGCAGAACAAAGTCCTGATTTCTGGGAATGTGTACAGCTAGCTTCTGT 81
Db 201 AAGAGTCTTCGAATCGGATTCGGGTTCTGGTGTTCCTCAATGTACTAGCTTGACTTAAT 260
QY 82 GAGAGGAGATTGGTTAGCAAGAGATACCATCATTTATGTCTCTCATTTGAGAGAGCCAA 141
Db 261 CAGAGATCGTTGATTAGCAAGAGACACCATCTCTTCAGCTTGATCGAGAGAGCTAA 320
QY 142 GTTCCCTAGCAATTTCTCACCTATGATGAAGAATGATCTCAATCCAGGGTTTTTGTGG 201
Db 321 GTTCCACTCAATTTCTCTGCTTTCGAGGAATCTCGTGTCTAGATTTCTGGAAGTTTCTC 380
QY 202 CTCATTTGGTGGAAATTTGTTTAAAGATACAGAGCCATTTCAAGCTAAGGCTGGAAGATA 261
Db 381 TTCTCTCACTGAGTTTTTCTCGAGAGACAGAAATCATCTCAAGCTAAGGTAGAGAGATA 440
QY 262 CAAAAACCCGTAAGAAAAAGCGCTTCTCCAGAAAAATTTACCAACATCAATTTGCCATC 321
Db 441 TGAATACCCGGAAGAGAAATCCCTTTCTCTTGAGAAACATCTCTCACTCGGTTTTCTCTAC 500
QY 322 TTACTCTCTTCAAACAGTTTTTTCATCTGCTGCTGCTCTCAATTAACATAAACAAGTCCAT 381
Db 501 GCACAAATATCATCGGCTTTGCAACCTTAAGCCTTATCTGTAAACATTAACAAACAAT 560
QY 382 CTGGAATAATGTAATTTCAAAGAGTTACTTCCATTTGCTGCTACTTCGGGTGATGATGGCAA 441
Db 561 CTGGGATATTTACTTTAAGAAATTCCTTCTTCTCAACCTGCGGATGATGGCAA 620
QY 442 CTATGGCAAACTGCGAGCTAATGACCTTTCAATTTGAGTCCATCTCTAGAGGATTCGA 501
Db 621 CTATCCATCACTGCTGCTAGTATCTGCGCTGTTTACAAGCTCTTTTCGAGAAGGATTCGA 680
QY 502 CTATGGAAGTTTGTAGCTGAGTGCAATTTCAAGGATGCTCTCAAGACTACGAGCCCTTT 561
Db 681 CTACGGTAATTTGTAGCTGAGTCAATTTAGAGATGCTCCACAAGATTAACAGCCTGC 740
QY 562 AATTCGAGCTAAGGATAAAGAGGATTTGATGAATTTGTTGACATTTTACAAGCGTTGAAGA 621
Db 741 GATTCGCGCTCAGGATAGAGAGGCTTTGATGAAGCTTTGACGTTTGAGAAAGTAGAAGA 800
QY 622 GACGTTGAGGAGAGAGTTGAAGAAGGCTGTGTTGTTGGCAGGAGTGAATCTTAA 681
Db 801 AATGTTTAAGAGAGAGTGCGAGAAGAAGCAAGAACGTTTGACAAGAAGTAAATTCAA 860
QY 682 -----CAGTGATGACATGACATGAACAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735
Db 861 CTCTGCTATCGCATGAGAGTGAAGAAGATATAAGTGGATCAATTTGCTTGCCTCTCG 920
QY 736 CTGTGACAAATTTGGGTGATACCTCTCACCAGAGGTTTCAAGTTGAGTACCTCTTGG 795
Db 921 CATCTACGGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTTCTACG 980
QY 796 CCGTCTAGACTGA 808
Db 981 TCGTCTCGATTGA 993

RESULT 13
L47356
LOCUS
DEFINITION Lycopersicon esculentum chorismate mutase mRNA, complete cds.
L47356
PLN 13-AUG-1999
```

```

ACCESSION L47356
VERSION L47356.1 GI:5732017
KEYWORDS chormate mutase; mutase.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 1000)
AUTHORS Eberhard, J., Bischoff, M., Raesecke, H.R., Amrhein, N. and Schmid, J.
TITLE Isolation of a cDNA from tomato coding for an unregulated,
cytosolic chorismate mutase
JOURNAL Plant Mol. Biol. 31 (4), 917-922 (1996)
MEDLINE 9640046
PUBMED 8806422
COMMENT GSDB:S:46596
FEATURES
     source
         1..1000
             Location/Qualifiers
                 /organism="Lycopersicon esculentum"
                 /mol_type="mRNA"
                 /strain="UC82b"
                 /db_xref="taxon:4081"
                 /clone="cLeCM2"
                 /EC_number="5.4.99.5"
                 /note="putative"
                 /codon_start=1
                 /product="chorismate mutase"
                 /protein_id="AAD48923.1"
                 /db_xref="GI:5732018"
                 /translation="WACQADKLCDSIRKSLIROETIIFNLIERIKFPINPTLYKQ
                 OLPPSNRSGSLFOYLFQTESLQKVGRLAPEENPPFDNLSIIPKTKTPVLH
                 PRAESVYVNEKILDIYNOMLPLECTEVNDANEATTAACDIQLLQALSRRIHYGKFV
                 AEVKRDSIDEXKPFILAQDRDALMKLITFEAVEWVKRVAKAKVFGQEVSLDNA
                 BEVKGIDPLLSRLYDEWVPLTKLVEVEYLRLRD"
     polyA_site
         1000
             /note="putative"
ORIGIN
     Query Match 29.2%; Score 297.8; DB 8; Length 1000;
     Best Local Similarity 64.5%; Pred. No. 3.1e-59; Indels 6; Gaps 2;
     Matches 478; Conservative 0; Mismatches 257;
QY 71 CTAGCTTCTGTGAGAGAGATTGGTTAGGCAAGAGATACCATTATTATGGTCTCATT 130
DB 107 CTTGATTCAATAAGAAAGATTTAATTAGACAAGAGATACCATTATTTCACCTTATT 166
QY 131 GAGAGAGCAAGTTCCTAGCAATTCACACTATGATGAAGATGCTCAATCCAG 190
DB 167 GAAAGAAATCAAAATTCCTCAATTAATCAATTTATACAAACAAATACCTCTCTCC 226
QY 191 GGTGTTTGGCTCATGTGGAAATTTGTTGTTAAGAATACAGAGCCATTCAAGCTAAG 250
DB 227 AATTTTTCAGATCCTTATTTCATCTTGTTCAGAAACAGATCCCTTCAATCTAAG 286
QY 251 GCTGGAAGATACAAAACCTGAGAAAAGCCCTTCCCGAGAAAATTTACCAACATCA 310
DB 287 GTTGGTAGATATTGGCCCCAGAGGAAAATCCTTTCTCCCGAGATAATTTGCTGACTCA 346
QY 311 ATTGTGCTATCTTACTTCTTCAAAAGTTTGTGATCCTGGTGGCTTCAATTAACATA 370
DB 347 ATCATACCACTTACAAATGCACACAGTTTGGATCTTGGAGAGATCAGTAATGTA 406
QY 371 AACAAATCCATCTGGAATAATGTAATTCAAAGAGTTACTTCCATTGGTTGCTACTCGG-- 428
DB 407 AATGAGAAAATATTGGATATTTATATAAATCAGATGCTTCCACTATTTTGTACCGAGGT 466
QY 429 -GTGATGATGGCAACTATGCGCAACTGCGAGCTAATGACCTTTTCATTATTGAGTCCATC 487
DB 467 AATGATGATGCAAAATTTTGGCACTACTGCTGCTTTGACATTCAGCTATTGAGGCACTA 526
QY 488 TCTAGAGAGATTCACATATGAAAGTTTGTAGCTGAGGTGAATTCAGGATGCTCCTCAA 547

```

---

```

Db 527 TCTAGAAAGATTCACTATGGAATAATTTGTTGCTGAAGTTAAATTTAGATAGATTGAT 586
QY 548 GACTTACGAGCCTTTAATTCGAGCTAAGGATAAAGAGGATTGATGAAATTTGTTGACATTT 607
Db 587 GAATATAAACCAATTTATTCTTCTCAGGATAGAGATGCTTTAATGAAGTTATTAAACATTT 646
QY 608 ACAAGCGTTGAAGAGAGCGGTGAGAGAGAGTTGAAAAGAGGCTGTGCTGTTTGGCGAG 667
Db 647 GAAGCGGTTTGAAGAGATGGTGAAGAGAGAGATAGCGAAGAGCTAAAGTGTGTTGCTCA 706
QY 668 GAAAGTGAATCTTAACAGTGTATGACAAATGACAAATGAAAACCGTAAATTTGATCCATCAGTG 727
Db 707 GAAAGTGAATCTTAAC--GATAATGCTGAAGAGTGAAGGCAAAATTCATCCATTTGTA 763
QY 728 GCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAGAGGAGGTTGAGTGAATAC 787
Db 764 GTTTCACGTTTATATGATGAATGGTTATGCTCTTGTCTGAACTTGTGCAAGTTGAATAC 823
QY 788 CTCTTGGCGCGCTCTAGACTGA 808
Db 824 CTCTGCTGCTCTTGTATTA 844

```

---

```

RESULT 14
AR404687 1231 bp DNA linear PAT 18-DEC-2003
LOCUS AR404687
DEFINITION Sequence 13 from patent US 6627798.
ACCESSION AR404687
VERSION AR404687.1 GI:40153355
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE 1 (bases 1 to 1231)
AUTHORS Falco, S.C., Famodu, O.O. and Lee, J.-M.
TITLE Aromatic amino acid biosynthetic enzymes
JOURNAL Patent: US 6627798-A 13 30-SEP-2003;
FEATURES
     Location/Qualifiers
         1..1231
             /organism="unknown"
             /mol_type="genomic DNA"
ORIGIN
     Query Match 25.2%; Score 257.4; DB 6; Length 1231;
     Best Local Similarity 58.4%; Pred. No. 8.9e-50;
     Matches 469; Conservative 0; Mismatches 331; Indels 3; Gaps 1;
QY 24 TGCCAAAGCAGCAGCAACAAAGTCTGATTCTCGGAATGTGTACACGCTAGCTTCTGTGA 83
DB 239 TGACCAAGGAGGAGAAACAGAGATAGATCAAAAGTGAATCTGACCTTGGACCAATTA 298
QY 84 GAGAGGATTTGGTTAGCAAGAGATACCATTATTATGGTCTCATTTGAGAGAGCCAGT 143
DB 299 GAACCTCCTTGGTTAGCAAGAGACAGCATCATATTTCAGCCTCTTTAGAGAGACAGT 358
QY 144 TCCTAGCAATTTCTCACCTATGATGAAAGTATGCTCAAAATCCAGGGTTTTGTGGCT 203
DB 359 TTGCTACAATGCTGATATATATGATAAATAATGCTTTCAATGGGATGATTTGATGGCT 418
QY 204 CATTTGGTGAATTTGTTTGAAGATACAGAGGCCATTCAAGCTAAGGCTGGAGAGATACA 263
DB 419 CTTTGGTTGAATTCATGTTTAGAGAAACCGAAAACCTACATCAACAGGTTGGAGATACA 478
QY 264 AAAACCTGAGAAAGCGCTTCTCCGAGAAATTTACCACCATCAATTTGTGCCATCTT 323
DB 479 AGAGCCCTGATAGACCCCATTTCTTCGGAGAGATCTGCTGNACCCTGTGTGCCACCTC 538
QY 324 ACTCCTTCAACAGATTTTTCATCCTGCTGGTCTGCTCTTCAATTAACATAAACAAGTCCATCT 383
DB 539 TCCAGTATCCAAAGGTTTTCATCCTTATTGCTGATTTCTATTATAATCAACAAGAGATTT 598
QY 384 GGAAATGATTTCAAGAGATTACTTCCATTGCTTCTGCTTCTCGGTGATGATGGCAACT 443

```





Db 365 TTGCTACAAATGCTGATATATATGATAAAATGCTTTCCATGTGATGGATTTGATGGCT 424  
Qy 204 CATTGTGGAATTTGTTGTAAGATACAGAGGCCATTCAAGCTAAGCTGGAAGATACA 263  
Db 425 CTTTGGTTGAATTCATGTTAGAGAAACCGAAAACTACATCAACAGGTTGGAGATACA 484  
Qy 264 AAAACCTGAGAAAGCGCTTCTCCAGAAAAATTTACCAACATCAATTTGCCATCTT 323  
Db 485 AGAGCCCTGATGAGCAACCCATTCTTTCCGAGGATCTGCCTGAACACATGTTGCCACCTC 544  
Qy 324 ACTCCTTCAAAAGTCTTTTGCACTCTGCTGCTCTCAATTAACATATAAACAAGTCCATCT 383  
Db 545 TCCAGTATCCAAAGTTTTCATCCTATTGCTGATCTATTATATCAACAAGGAGATTT 604  
Qy 384 GGAATATGTAATTCAAAGAGTTACTTCCATTTGCTGCTACTTCCGGGTGATGAGCAACT 443  
Db 605 GGAATATGTAATTTGATGAGCTTCTTCCAAGATTAGTGAAGAAGGAGTGTGTTAATT 664  
Qy 444 ATGGCAAACTGCAGCTAATGACCTTTTCATTATTGCACTCCATCTCTAGAAGGATTCCT 503  
Db 665 ATGATCCAGTGTCTTTTGACACAGATCTGTTGAGGGCTCTCCAAAGAAATTCCT 724  
Qy 504 ATGGAAGTTTGTAGCTGAGGTGAAATTCAGGAGTCTCTCAAGACTACGAGCCTTTAA 563  
Db 725 ATGTAAGTTTGTGGCAGAGGCTAAGTTTCAAGAGTCTCTGAAGCTTACATGCCTGCGA 784  
Qy 564 TTCGAGCTAAGGATAAAGAGGATTGATGAATTTGACATTTACAGCGTTGAAGAGA 623  
Db 785 TAATAGCACAGGACTGCGATCAACTAATGACCTCTCCTACCTATGAACCGTGGAGCGTG 844  
Qy 624 CGGTGAGGAAGAGAGTTGAAAGAGGCTGTGTTGTTGGCAGGAGTGAATCTTAACA 683  
Db 845 CTATTGAACATAGGTGGAAGCTAAGCTAAGATCTTTGGACAGGAGTGGATTTAGGCG 904  
Qy 684 GTATGACATGACAAT---GAAACCGTAATTTGATCCATCAGTGGCTTCTAGCTTGT 740  
Db 905 CTGAAGACAAACGGCGCTCCACCAATGTACAGATAAGGCCCAAGTTTGGTGGCTGAACGT 964  
Qy 741 ACAAAATTTGGTGATACCTCTCACAAGGAGGTTGAGTTGAGTACCTCTTTCGCCCTC 800  
Db 965 ACAGCTACAGGATCATGCCGCTAACCAAGGAGGTTCAAGTAGCCTACTTGTGAGGAGAT 1024  
Qy 801 TAGACTGAGGCAATTACAATGCA 823  
Db 1025 TGGATTGATTGTTTACGATTGTA 1047

Search completed: September 25, 2004, 00:33:08  
Job time : 4169 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 22:28:41 ; Search time 477 Seconds  
(without alignments)  
9084.197 Million cell updates/sec

Title: US-10-624-061-15  
Perfect score: 1020  
Sequence: 1 gcacgagagcttaattcaa.....atcttcaaaaaaaaaa 1020

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*  
1: Geneseq1930s.\*  
2: Geneseq1930s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001as.\*  
5: Geneseq2001bs.\*  
6: Geneseq2002s.\*  
7: Geneseq2003as.\*  
8: Geneseq2003bs.\*  
9: Geneseq2003cs.\*  
10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020	100.0	1020	7	ABX93058 Soybean c
2	485.6	45.6	525	7	ABX93053 Soybean c
3	327.2	32.1	1006	6	ABA99613 A. thalia
4	327.2	32.1	1006	6	ABA91400 Arabidops
5	327.2	32.1	1006	6	ABA91394 Arabidops
6	326.2	32.0	798	6	ABZ13167 Arabidops
7	326.2	32.0	993	6	ABA99614 Construct
8	257.4	25.2	1231	7	ABX93057 Rice chor
9	255	25.0	1274	3	AAC33324 Arabidops
10	254.8	25.0	1207	6	ABA91393 Arabidops
11	254.8	25.0	1207	6	ABA91399 Arabidops
12	251	24.6	1005	6	ABZ12932 Arabidops
13	239.4	23.5	1223	7	ABX93056 Corn chor
14	229.4	22.5	780	7	ABX93059 Wheat cho
15	211.4	20.7	1217	6	ABA91395 Arabidops
16	211.4	20.7	1217	6	ABA91401 Arabidops
17	206.2	20.2	579	7	ABX93054 Wheat cho
18	99.8	9.8	843	3	AAC81949 H. polymo
19	99.8	9.8	1655	3	AAC81950 H. polymo
20	98.8	9.7	5176	6	ABA91397 Arabidops
21	93.8	9.2	258	6	ABL71526 Corn tass
22	84	8.2	287	6	ABL72637 Corn tass
23	73.6	7.2	5099	6	AAD29643 Arabidops

24	73.6	7.2	5099	6	ABA91396 Arabidops
25	69.2	6.8	552	7	ABX93051 Corn chor
26	58.4	5.7	601	7	ABX93052 Rice chor
27	53.8	5.3	869	7	AAL50202 M sterili
28	52.2	5.1	658	3	AAF14469 Aspergill
29	52.2	5.1	2920	7	ABT17891 Aspergill
30	52.2	5.1	2987	7	ABT19705 Aspergill
31	51.4	5.0	560	4	AAH44222 Physcomit
32	47.6	4.7	2000	7	ADA71938 Rice gene
33	44	4.3	4149	6	ABK97892 DNA encod
34	43.4	4.3	2000	7	ADA71938 Rice gene
35	42.4	4.2	113515	6	ABL34175 Human inm
36	42	4.1	5379	6	ABL33676 Human inm
37	42	4.1	5379	6	ABL34576 Human met
38	42	4.1	5379	6	ABL70369 Chemical
39	40.8	4.0	6219	6	ABL32866 Human inm
40	40.8	4.0	6219	6	AAS63324 Chemical
41	40.4	4.0	5311	6	ABL33018 Human inm
42	40.4	4.0	7455	6	ABL33759 Human inm
43	39.8	3.9	6418	6	ABL32323 Human inm
44	39.8	3.9	6418	6	AAS61074 Human gen
45	39.8	3.9	110000	2	AAK20248_08 Continuation (9 of

## ALIGNMENTS

RESULT 1  
ABX93058  
ID ABX93058 standard; cDNA; 1020 BP.  
XX  
AC ABX93058;  
DT 22-MAY-2003 (first entry)  
XX  
DE Soybean chorismate mutase cDNA.  
XX  
KW Soybean; plant; gene; ss; chorismate mutase; aromatic amino acid; diet;  
KW biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;  
KW phenylalanine; prephenate dehydrogenase; herbicide; food crop.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 23..808  
FT /\*tag= a  
FT /product= "Chorismate mutase"  
XX  
PN US2002184658-A1.  
XX  
PD 05-DEC-2002.  
XX  
PF 03-DEC-1999; 99US-00454279.  
XX  
PR 04-DEC-1998; 98US-0110845P.  
XX  
PA (FALC/) FALCO S C.  
PA (FAMO/) FAMODU O O.  
PA (LEEG/) LEE J.  
XX  
DR Falco SC, Famodu OO, Lee J;  
XX  
DR WPI: 2003-328651/31.  
XX  
DR P-PSDB; ABU08094.  
XX  
PT New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
PT nucleic acids, useful for facilitating design and/or identifying  
PT inhibitors of those enzymes that may be used as herbicides and for  
PT producing antibodies.  
XX  
PS Claim 3; Page 22; 32pp; English.  
XX  
CC The invention discloses isolated polynucleotides encoding chorismate

mutase polypeptides. Aromatic amino acids must be included in the diets of animals. In the aromatic amino acid biosynthetic pathway chorismate is converted to anthranilate during tryptophan biosynthesis and is converted to prephenate, the branch point for tyrosine and phenylalanine biosynthesis. Chorismate mutase catalyses the conversion of chorismate to prephenate. Also disclosed are methods for selection of an isolated polynucleotide that affects the level of expression of a tyrosine biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid fragment encoding a tyrosine biosynthetic enzyme polypeptide and evaluating compounds for their ability to inhibit the activity of a tyrosine biosynthetic enzyme. The polypeptides can be used to produce antibodies. Chorismate mutase and prephenate dehydrogenase are good targets for herbicides that will not affect animals, and overexpression of these enzymes may be used to increase the content of aromatic amino acid in food crops. The polypeptides may also be used to design and/or identify inhibitors of those enzymes that may be used as herbicides. The nucleic acids may be used to create transgenic plants, as probes for the genetic and physical mapping of the genes and as markers for traits linked to those genes. The sequence presented is a soybean chorismate mutase cDNA

XX Sequence 1020 BP; 322 A; 176 C; 223 G; 299 T; 0 U; 0 Other;

Query Match 100.0%; Score 1020; DB 7; Length 1020;  
Best Local Similarity 100.0%; Pred. No. 5.9e-268;  
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGAGCTTAATTCAGAGATGCCAAAGCAGCAGAACAAAGCTCTGGGAA 60  
DB 1 GCACGAGAGCTTAATTCAGAGATGCCAAAGCAGCAGAACAAAGCTCTGGGAA 60  
QY 61 TGTGTACACGCTAGCTCTCTGAGAGAGGATTTGGTTAGCAAGGATACCATCATTTA 120  
DB 61 TGTGTACACGCTAGCTCTCTGAGAGAGGATTTGGTTAGCAAGGATACCATCATTTA 120  
QY 121 TGGTCTCATTCAGAGAGCCAGTTCCTAGCAATTCACACCTAGTGAAGAGTATGC 180  
DB 121 TGGTCTCATTCAGAGAGCCAGTTCCTAGCAATTCACACCTAGTGAAGAGTATGC 180  
QY 181 TCAATCCAGGGTTTTGTGGCTCAITGGTGAATTTGTTTAAAGATACAGAGCCAT 240  
DB 181 TCAATCCAGGGTTTTGTGGCTCAITGGTGAATTTGTTTAAAGATACAGAGCCAT 240  
QY 241 TCAAGCTAAGCGTGGAGATACAAAACCTCGAAGAAACGCTTCTCCAGAAATTT 300  
DB 241 TCAAGCTAAGCGTGGAGATACAAAACCTCGAAGAAACGCTTCTCCAGAAATTT 300  
QY 301 ACCACCATCAATTTGTCATCTTACTCCTTCAACACAGTTTTTGCTGCTGCTTC 360  
DB 301 ACCACCATCAATTTGTCATCTTACTCCTTCAACACAGTTTTTGCTGCTGCTTC 360  
QY 361 AATTAAACATAAACAAGTCCATCTGGAATATGTAATTTCAAGAGTTACTCCATTCCTGC 420  
DB 361 AATTAAACATAAACAAGTCCATCTGGAATATGTAATTTCAAGAGTTACTCCATTCCTGC 420  
QY 421 TACTTCGGGTGATGCGCACTATGCGCAATCTGAGCTTAATGACCTTTCATTATTGCA 480  
DB 421 TACTTCGGGTGATGCGCACTATGCGCAATCTGAGCTTAATGACCTTTCATTATTGCA 480  
QY 481 GTCCATCTCTAGAGGATTCATATGGAAGATTTGTTAGCTGAGGTGAATTCAGGATGC 540  
DB 481 GTCCATCTCTAGAGGATTCATATGGAAGATTTGTTAGCTGAGGTGAATTCAGGATGC 540  
QY 541 TCCTCAAGACTACGAGCTTTAATTCAGCTAAGGATTAAGAGGATTCATGAATTTGT 600  
DB 541 TCCTCAAGACTACGAGCTTTAATTCAGCTAAGGATTAAGAGGATTCATGAATTTGT 600  
QY 601 GACATTTACAAGCGTTGAAGAGAGCGGTGAGGAAGAGGTTGAAAAGAGCGTGTGTGT 660  
DB 601 GACATTTACAAGCGTTGAAGAGAGCGGTGAGGAAGAGGTTGAAAAGAGCGTGTGTGT 660  
QY 661 TGGCAGAGAGTGAATCTTAACAGTGATGACATGACATGAATGAACCGTAAATTCATCC 720  
DB 661 TGGCAGAGAGTGAATCTTAACAGTGATGACATGACATGAATGAACCGTAAATTCATCC 720

661 TGGCAGAGAGTGAATCTTAACAGTGATGACATGAATGAACCGTAAATTCATCC 720  
721 ATCAGTGGCTTCTAGCTTGTACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGT 780  
721 ATCAGTGGCTTCTAGCTTGTACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGT 780  
781 TGAGTACCTCTTGGCGGCTCTAGACTGAAGGCATTACAATGCAAGTTAGAAATTTAGAGAA 840  
781 TGAGTACCTCTTGGCGGCTCTAGACTGAAGGCATTACAATGCAAGTTAGAAATTTAGAGAA 840  
841 TGGAGATGAATATGATGTTGTTTCAATGATTAAGCTCTTAAGTGATCCTTTATTTC 900  
841 TGGAGATGAATATGATGTTGTTTCAATGATTAAGCTCTTAAGTGATCCTTTATTTC 900  
901 CAATCTCATCTAGCTGTTGATTCAGAAATATTTTGTAGCTATAGATTGTTACCTTTAT 960  
901 CAATCTCATCTAGCTGTTGATTCAGAAATATTTTGTAGCTATAGATTGTTACCTTTAT 960  
961 TTTTCATCGGCTTTATTAGGAAAAGGCATTATTATATCATGATCTTCAAAAAA 1020  
961 TTTTCATCGGCTTTATTAGGAAAAGGCATTATTATATCATGATCTTCAAAAAA 1020

RESULT 2  
ABX93053  
ID ABX93053 standard; cDNA; 525 BP.  
XX  
AC ABX93053;  
XX  
DT 22-MAY-2003 (first entry)  
XX  
DE Soybean chorismate mutase expressed sequence tag.  
XX  
KW Soybean; plant; gene; ss; chorismate mutase; expressed sequence tag; EST;  
KW aromatic amino acid; diet; biosynthesis; chorismate; anthranilate;  
KW tryptophan; prephenate; tyrosine; phenylalanine;  
KW prephenate dehydrogenase; herbicide; food crop.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 43..438  
FT /\*tag= a  
FT /product= "Chorismate mutase"  
FT /partial  
FT /note= "No start or stop codon shown"  
XX  
PN US2002184658-A1.  
XX  
PD 05-DEC-2002.  
XX  
PF 03-DEC-1999; 99US-00454279.  
XX  
PR 04-DEC-1998; 98US-0110845P.  
XX  
PA (FALC/) FALCO S C.  
PA (FAMO/) FAMODU O O.  
PA (LEEJ/) LEE J.  
XX  
PI Falco SC, Famodu OO, Lee J;  
XX  
XX WPI; 2003-328651/31.  
DR P-PSDB; ABU08089.  
XX  
PT New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
PT nucleic acids, useful for facilitating design and/or identifying  
PT inhibitors of those enzymes that may be used as herbicides and for  
PT producing antibodies.  
XX  
PS Claim 3; Page 16; 32pp; English.  
XX  
CC The invention discloses isolated polynucleotides encoding chorismate  
mutase polypeptides. Aromatic amino acids must be included in the diets

of animals. In the aromatic amino acid biosynthetic pathway chorismate is converted to anthranilate during tryptophan biosynthesis and is converted to prephenate, the branch point for tyrosine and phenylalanine biosynthesis. Chorismate mutase catalyses the conversion of chorismate to prephenate. Also disclosed are methods for selecting an isolated polynucleotide that affects the level of expression of a tyrosine biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid fragment encoding a tyrosine biosynthetic enzyme polypeptide and evaluating compounds for their ability to inhibit the activity of a tyrosine biosynthetic enzyme. The polypeptides can be used to produce antibodies. Chorismate mutase and prephenate dehydrogenase are good targets for herbicides that will not affect animals, and overexpression of these enzymes may be used to increase the content of aromatic amino acid in food crops. The polypeptides may also be used to design and/or identify inhibitors of those enzymes that may be used as herbicides. The nucleic acids may be used to create transgenic plants, as probes for the genetic and physical mapping of the genes and as markers for traits linked to those genes. The sequence presented is a soybean chorismate mutase expressed sequence tag (EST)

Sequence 525 BP; 160 A; 103 C; 111 G; 148 T; 0 U; 3 Other;

Query Match 45.6%; Score 465.6; DB 7; Length 525;  
Best Local Similarity 95.4%; Pred. No. 9.5e-117;  
Matches 499; Conservative 0; Mismatches 22; Indels 2; Gaps 2;

QY 8 AGCTTAATTCAGAAATGCCAAGCAGCAGCAAAAGTCTGTATCTGGGAATGTGTAC 67  
DB 1 AGCTTAATTCAGAAATGCCAAGCAGCAGCAAAAGTCTGTATCTGGGAATGTGTAC 60  
QY 68 ACGTGTAGCTTCTGTGAGAGAGATTTGTTAGGCAAGAGATACCATTTATGGTCTC 127  
DB 61 ACGTGTAGCTTCTGTGAGAGAGATTTGTTAGGCAAGAGATACCATTTATGGTCTC 120  
QY 128 ATTGAGAGAGCAAGTTCCTTAGCAATTCACACCTATGATGAAAGATGTCTCAATC 187  
DB 121 ATTGAGAGAGCAAGTTCCTTAGCAATTCACACCTATGATGAAAGATGTCTCAATC 180  
QY 188 CAGGGTTTTGTGGCTCATTTGGTGAATTTGTTTGAATATACAGAGCCATTCAGCT 247  
DB 181 CAGGGTTTTGTGGCTCATTTGGTGAATTTGTTTGAATATACAGAGCCATTCAGCT 240  
QY 248 AAGGCTGAAGATACAAAACCCCTGAAGAAACCGCTTCTTCCAGAAAATTTACACCA 307  
DB 241 AAGGCTGAAGATACAAAACCCCTGAAGAAACCGCTTCTTCCAGAAAATTTACACCA 300  
QY 308 TCAATTTGCCATCTTACTCTTCAACAGATTTTGGATCTGTGCTGTCTCAATTAAC 367  
DB 301 TCAATTTGCCATCTTACTCTTCAACAGATTTTGGATCTGTGCTGTCTCAATTAAC 360  
QY 368 ATAAACCAAGTCCATCTGGAAATGATTTTCAAGAGTACTTCCATTTGCTTACTTCG 427  
DB 361 ATAAACCAAGTCCATCTGGAAATGATTTTCAAGAGTACTTCCATTTGCTTACTTCG 420  
QY 428 GGTGATGATGCAATATGCGCAAACTGCGACTAATGACCTTCAATTTATGATGCAATC 487  
DB 421 GGTGATGATGCAATATGCGCAAACTGCGACTAATGACCTTCAATTTATGCAATC 479  
QY 488 TCTAGAGGATCTACTATGGAATCTTCTAGCTCAGGTGAAT 530  
DB 480 TCTAGAGGA-TCACTATGGAAGATTTGTAGCTGANGGAAT 521

RESULT 3

ABA99613

XX ABA99613 standard; DNA; 1006 BP.

AC ABA99613;

XX ABA99613;

XX 17-MAY-2002 (first entry)

XX A. thaliana cytosolic chorismate mutase-2 DNA.

XX

KW Shikimate pathway; chorismate mutase; gene; prephenate dehydrogenase;  
KW vitamin E; vitamin K; ubiquinone; fat-soluble antioxidant; transgenic;  
KW feed; pharmaceutical; cosmetic; aromatic amino acid; salicylic acid;  
KW folic acid; phenylpropanoid; flavanoid; plant; abiotic stress resistance;  
KW tocopherol; tocotrienol; ds.  
XX Arabidopsis thaliana.

XX Key Location/Qualifiers  
XX CDS 64..861  
XX FT /\*tag= a  
XX FT /product= "chorismate mutase-2"

XX WO200200901-A1.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-EP007391.

XX 29-JUN-2000; 2000DE-01030647.

XX 21-DEC-2000; 2000DE-01064454.

XX (SUNG-) SUNGENE GMBH & CO KGAA.

XX Badur R, Geiger M, Kunze I, Sommer S;

XX WPI; 2002-164442/21.

XX P-PSDB; AAM49654.

XX Preparing fine chemicals, particularly Vitamins E and K, useful as  
XX antioxidants e.g. in foods or medicine, by growing organisms with altered  
XX shikimate biosynthesis pathway.  
XX Disclosure; Page 57-60; 74pp; German.

XX This invention describes a novel method for the preparation of fine  
XX chemicals by culturing organisms in which the shikimate pathway has been  
XX altered relative to the wild type. The method involves a construct  
XX containing a plasmid transit peptide, a chorismate mutase and/or  
XX prephenate dehydrogenase, linked to at least one regulatory sequence for  
XX transcription or translation in plants. The method is used to produce  
XX chemicals, particularly Vitamins E or K and/or ubiquinone, useful e.g. as  
XX fat-soluble antioxidants (in feeds, pharmaceuticals and cosmetics), but  
XX also aromatic amino acids, salicylic or folic acid derivatives.  
XX phenylpropanoids, flavanoids etc., especially in transgenic plants. Also  
XX plants with increased Vitamin E contents have improved resistance to  
XX abiotic stress, e.g. frost and drought. Transgenic plants with an altered  
XX shikimate pathway are useful as foods, fodder and in preparation of  
XX processed foodstuffs. Transgenic plants with modified shikimate pathways  
XX have increased content of desired chemicals, particularly tocopherols  
XX and/or tocotrienols. This sequence encodes the Arabidopsis thaliana  
XX cytosolic chorismate mutase-2 protein used in the method of the invention

XX Sequence 1006 BP; 290 A; 188 C; 226 G; 302 T; 0 U; 0 Other;

Query Match 32.1%; Score 327.2; DB 6; Length 1006;

Best Local Similarity 64.2%; Pred. No. 7.4e-79;

Matches 510; Conservative 0; Mismatches 278; Indels 5; Gaps 1;

QY 22 ATGCGCAAGCAGCAGCAAAAGTCTGTATCTTGGGAATGTGTACGCTTCTGT 81  
DB 69 AAGAGTCTTCGAATCGGATTCGGTTCTGTGTTTCCAAATGTACTGTGACTTAAAT 128  
QY 82 GAGAGAGGATTTGGTTAGCAGAGGATACCATCTTTATGTCTCATTTGAGAGGCCAA 141  
DB 129 CAGAGAAATCGTTGATTTAGCAGAGAGACACCATCTCTTCAGCTTGTTCGAGAGCTAA 188  
QY 142 GTTCCTAGCAATCTCACACCTATGATGAAAGATGCTCAAATCCAGGGTTTTTGG 201  
DB 189 GTTCCACTCAATCTCTCTGCTTTCCAGGAATCTCGTGTCTAGATTCTGGAAGTTCTC 248  
QY 202 CTCATTGGTGGATTTGTTTGAATACAGAGCCATTCAAGCTTAGCTGGAGATA 261



QY	736	CTTGTACAAAATTTGGTGTATACCTCTCACCAAGGAGGTTTCAGGTTGAGTACCTCTTGGG	795		
DB	218	CATCTACGGGGAAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTGAGTATCTTCTACG	159		
QY	796	CCGCTCTAGACTGAA	809		
DB	158	TCGCTCTCGATTGAA	145		
RESULT 5					
ABA91394					
ID	ABA91394	standard; cDNA; 1006 BP.			
XX	AC	ABA91394;			
DT	08-APR-2002	(first entry)			
DE	Arabidopsis chorismate mutase-2 cDNA.				
XX	KW	Chorismate mutase-2; herbicide; plant; enzyme; gene; ss.			
OS	Arabidopsis thaliana.				
XX	PN	WO200202798-A2.			
XX	PD	10-JAN-2002.			
XX	PF	22-JUN-2001; 2001WO-US020104.			
XX	PR	05-JUL-2000; 2000US-00610040.			
XX	PA	(PARA-) PARADIGM GENETICS INC.			
PI	Boyes DC, Davis KR, Woessner JP, Goriach J, Hamilton CM;				
PI	Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;				
PI	Kjemtrup S;				
XX	WPI; 2002-154754/20.				
DR	Identifying compounds useful for modulating and inhibiting plant growth				
XX	and development, by using chorismate synthase and chorismate mutase as				
PT	targets which are essential for plant growth.				
XX	Claim 37; Page 64-65; 73pp; English.				
XX	The present sequence is that of a full-length cDNA encoding Arabidopsis				
CC	thaliana chorismate mutase-2 (CM), an enzyme that catalyses the				
CC	conversion of chorismate to prephenate in a branch of the shikimate				
CC	pathway leading to phenylalanine and tyrosine biosynthesis. CM is				
CC	essential for plant growth; inhibition of CM gene expression in plant				
CC	seedlings results in severe chlorosis, reduced growth and developmental				
CC	abnormalities. The invention provides compositions for the modulation of				
CC	plant growth or development using CM antisense (see ABA91400) and sense				
CC	polynucleotides, double-stranded RNA and ribozymes, and related				
CC	expression cassettes and vectors. The compositions are particularly				
CC	useful for the modulation and inhibition of plant growth, and for				
CC	generating male sterile plants or plants that produce seedless fruit. CM				
CC	is also a target for the identification of herbicides, and methods are				
CC	provided for using CM in the identification of herbicide candidates.				
CC	Methods are also provided for identifying compounds that stimulate the				
CC	expression or function of CM. These can be used to promote growth and				
CC	development				
XX	SQ Sequence 1006 BP; 290 A; 188 C; 226 G; 302 T; 0 U; 0 Other;				
Query Match					
Best Local Similarity 32.1%; Score 327.2; DB 6; Length 1006;					
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;					
QY	22	AATGCCCAAGCAGCAGAACAAAGTCCTGATTCCTGGGAATGTCACACGCTAGCTTCTGT	81		
DB	69	AAGAGTCTTCGAATCCGATTCGGGTTCGGTTCCTCAATGACTGACTGCTGACTTAAT	128		

QY	82	GAGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTATTTGGTCTCTATTGAGAGACCCA	141
DB	129	CAGAGAATTCGTTGATTAGGCAAGAGACACCATCTGCTTCCAGCTTTCAGTTCAGAGAGCTAA	188
QY	142	GTTCCTAGCAATTTCTCACACCTATGATGAAAAGTATGCTCAAAATCCAGGGTTTTTGTGG	201
DB	189	GTTCCTCACTCAATTTCTCTGCTTCGAGGAACTCTCGTTGTCTAGATTTCTGGAAGTTTCTC	248
QY	202	CTCATTTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAGACTTAAGCTTGAAGATA	261
DB	249	TTCTCTCACTGAGTTTTTCGTCAGAGAGACAGAAATCATCCAAGCTTAAGGTAGGAAGATA	308
QY	262	CAAAAAACCTCAGAAAAACGGCTTCTTCCAGAAAAATTTACCACCACTCAATTTGCGCATC	321
DB	309	TGATACCCGGAAGAGAACTCTTTCCTTCTTCTTCTGAGACATTCCTCACTCGGTTTTCTTAC	368
QY	322	TTACTCCTTCAAAACAGTTTTTTCATCTCTGCTTCAATTTAAACATAAAACAAGTCCAT	381
DB	369	GCACAAATATCATTCGGCTTTGACACCCCTAAGGCTCTATCTGTGTTAAACATTAACAAACAAT	428
QY	382	CTGGAAATGTAATTTCAAGAGGTTACTTCCATTTGCTTGTCTACTTCCGGTGATGATGGCAA	441
DB	429	CTGGGATATTTACTTTTAAAGAAATGCTTCTTCTTGTGTTGTCAAACCTGGCGATGATGGCAA	488
QY	442	CTATCGCAAACTCGACCTAATGACCTTTTCATTTATTCAGTCCATCTCTAGAAGGATTC	501
DB	489	CTATCCATCAACTGCTGCTAGTGATCTCGCTGTTGTACAAAGCTCTTTCGAGAAGGATTC	548
QY	502	CTATGGAAGTTTCTAGCTGAGGTGAAATTCAGGAGTCTCTCTCAAGACTACGAGCCTTT	561
DB	549	CTACGGTAAATTTGAGCTGAGGTCAAATTCAGAGATGCTCCACAAAGATTACGAGCGCTG	608
QY	562	AATTCGAGCTAAGGATAAAGAGGATTTGATGAAATTTGTCACATTTTCAAGCGTTGAAGA	621
DB	609	GATTCGCGCTCAGGATAGAGAGGCTTTGATGAAGCTGTTGACGTTTGAGAAAAGTAGAAGA	668
QY	622	GACGTGAGGAAGAGAGTTGAAAGAGGCTGTGGTGTGTTGGCAGGAGTGAATCTTAA	681
DB	669	AATGGTTTAAAGAGAGGTGCGAAGAAAGCAGAAAACGTTTGGCAAGAAAGTAAAAATCAA	728
QY	682	-----CAGTGATGACAATGACAATGAAAACCGTAAATTTGATCCATCAGTGGCTTCTAG	735
DB	729	CTCTGGCTATGGCGATGAGATAGAGAAGATATAAGTGGATCCATTCCTTCCCTCTCG	788
QY	736	CTTGTACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGTTGAGTACCTCTTGGG	795
DB	789	CATCTACCGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTTCTACG	848
QY	796	CCGCTCTAGACTGAA	809
DB	849	TCGCTCTCGATTGAA	862
RESULT 6			
ID	ABZ13167	standard; DNA; 798 BP.	
XX	AC	ABZ13167;	
XX	DT	21-JAN-2003 (first entry)	
XX	DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 972.	
XX	KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO200216655-A2.	
XX	XX	28-FEB-2002.	
XX	XX	24-AUG-2001; 2001WO-US026895.	
XX	XX		

QY	82	GAGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTATGGTCTCATTTAGAGAGCCAA	141
DB	129	CAGAGATCGTTGATTTAGGCAAGAGAGACACCATCGTCTTCAGTTGATCGAGAGACTAA	188
QY	142	GTTCCCTAGCAATTCACACCATATGATGAAAGATATGCTCAATCCAGGTTTTTGTGG	201
DB	189	GTTTCACACTCAATTCCTCGCTTTCGAGGAATCTCGTTGTCTAGATTCTGGAAGTTTCTC	248
QY	202	CTCATTTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAAGCTAAAGCTTGGAGATA	261
DB	249	TTCTCTCACTGAGTTTTTCGTCAGAGACAGAAATCATCAAGCTAGGTAGTAGAGATA	308
QY	262	CAAAAACCCCTGAAGAAAAGCCCTTCTCCAGAAAATTTACACCATCAATTTGCCATC	321
DB	309	TGAATACCCGGGAAGAGAAATCCCTTCTCTGAGAACATTCCTCACTCGGTTTTTCTAC	368
QY	322	TTACTCTCTCAACACAGTTTTTTCATCCCTGCTGCTCTCAATTAACATAAACAAGTCCAT	381
DB	369	GCACAAATATCCATCGGCTTTTCACCCTAAGGCTCTATCTGTTAAATTAACAAACAAT	428
QY	382	CTGGAAATGTATTTCAAAGAGTTACTTCCATTTGCTTGTCTACTTCGGGTGATGGCAA	441
DB	429	CTGGGATATTTACTTTAAAGAAATTCCTTCTCTTTGTTGTCRAACCTGGCGATGATGGCAA	488
QY	442	CTATGGGCAACTGCAGCTAATGACCTTTCATTTATGTCAGTCCATCTCTAGAGGATTC	501
DB	489	CTATCCATCACTGCTGCTAGTATCTCGCTGTTTACAAGCTCTTTCGAGAGGATTC	548
QY	502	CTATGGAAGTTTGTAGCTGAGTGCAATTCAGGAGTCTCTCAAGACTACGAGCCCTTT	561
DB	549	CTAAGGTAAATTTGTAGCTGAGTCAATTCAGAGATGCTCCACAAGATTACGAGCCCTG	608
QY	562	AATTCAGCTAAGGATAAAGAGGATTGATGAAATTTGTTGACATTTACAGCGTTGARA	621
DB	609	GATTCGCGCTCAGGATAGAGAGGCTTTGATGAAGCTGTTGACGTTTGAGAAAGTAGA	668
QY	622	GACGGTGAAGAGAGATTGAAAGAGGCTCTGCTGTTTGGGCGAGGAAGTGAATCTTAA	681
DB	669	AATGTTAAGAGAGAGTCCAGAGAGAGAGAGAGGTTTGGACAGAGATTAATAATTCAA	728
QY	682	-----CAGTGATGACAAATGACAAATCAAAACCGTAAATTTGATCCATCAGTGGCTTAG	735
DB	729	CTCTGGCTATGCGATGAGAGTAAGAGAAAGTATTAAGTGGATCCATTGCTTGCCTCTCG	788
QY	736	CTTGTAACAAATTTGGTGATACCTCTCACCAGGAGGTTTCAGGTTGAGTACCTCTTGG	795
DB	789	CATCTACGGGAATGCTTATCCCTCTCACTAAGCTCGTTGAGGTGAGTATCTTCTACG	848
QY	796	CCGCTCTAGACTGAA	809
DB	849	TCGCTCGATTGAA	862
RESULT 6			
ABZ13167			
ID	ABZ13167	standard; DNA; 798 BP.	
XX	AC	ABZ13167;	
XX	AC	ABZ13167;	
DT	21-JAN-2003	(first entry)	
DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 972.		
XX	XX	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO200216655-A2.	
XX	PD	28-FEB-2002.	
XX	PF	24-AUG-2001; 2001WO-US026685.	
XX	XX		







Db 239 TGACCAAGGAGGAGAAACAGAGATAGATCAAAAGTGAAATACTGACCTTGGACAACTTA 298  
Qy 84 GAGAGATTGGTTAGGCAAGAGATACCATCATTTATGTTCTCATTTGAGAGAGCAAGT 143  
Db 299 GAACCTCTTGGTTAGGCAAGAGATACCATCATTTATGTTCTCATTTGAGAGAGCAAGT 358  
Qy 144 TCCTAGCAATTTCTCACCTATGATGATAAGATGATGCTCAAAATCCAGGGTTTTTGCGT 203  
Db 359 TTTGCTACAAATGCTGATATATATGATAAAATGCTTTCCATGCTGATGATTTGATGGT 418  
Qy 204 CATTGGTGAATTTGTTGTAAGATACAGAGGCCATTTCAAGCTAAGGCTGGAATACATA 263  
Db 419 CTTTGGTGAATTCATGGTTAGAGAACCGAAACTATACATCAACAGGTTGGAGATACA 478  
Qy 264 AAAACCTCTGAAGAAACGCTTCTTCCAGAAAATTTACCAATCAATTTGCGCATCTT 323  
Db 479 AGAGCCCTGATGAGCAACCCATCTTTCGGAGGATCTGCTGAACCACTGTTGCCACCTC 538  
Qy 324 ACTCCTTCAACAGATTTTTCATCTCTGCTGCTCTCAATTAACATAAACAAGTCACT 383  
Db 539 TCCAGTATCCAAAGGTTTTCATCTCTGCTGATTTCTGATTAATATCAACAAGGAGATT 598  
Qy 384 GGAATATGATTTCAAGAGATTTACTTCCATTTGCTTCTTCTCGGCTGATGATGCACT 443  
Db 599 GGAATATGATTTTGTATGAGCTTCTTCCAGATTTAGTGAAGAGGAAGTGAATTT 658  
Qy 444 ATCGCAAACTGCAAGTATGACCTTTCTATTTATGTCAGTCCATCTCTGAAGAGATTCACT 503  
Db 659 ATGATCCAGTCTCTTTGTGACACGATCTGTTGCGAGCGCTCTCCAAAAGAAATTCAT 718  
Qy 504 ATGGAAGTTTGTAGCTGAGTCAATTTCAAGATGCTCTCAAGACTACGAGCTTTTAA 563  
Db 719 ATGTAAGTTTGTGAGAGGCTTAAGTTTCAAGAGTCTCTGAGCTTACATGCTTGGCA 778  
Qy 564 TTCGAGCTAAGGATAAGAGAGGATTTGATGAAATTTGTCATTTAACAAGCGTTGAAGAGA 623  
Db 779 TAATAGCACAGGACTCGATCAACTAATGCACTCTCACTATGAACGGTGGAGCGTG 838  
Qy 624 CGTGAGGAGAGAGTTGAAGAGCGCTGCTGTTTGGCGAGGAGTGAATCTTACA 683  
Db 839 CTATTGAACATAGGTTGAAGCTTAAGCTTAAGCTTTTGGACAGAGGTTGATTTAGGG 898  
Qy 684 GTGATGACATGACAAAT---GAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTTGT 740  
Db 899 CTGAGACACAGCGGCTCCACCAATGTACAAGATAGGCCAGTTTGGTGGCTGAACGT 958  
Qy 741 ACAAAATTTGGTGTATACCTCTCCACAGGAGTTCAAGTTGAGTACCTTTGCGCGCTC 800  
Db 959 ACAGCTACAGGATCATGCGCTTAACAGAGGTTCAAGTAGCCTACTTGTGAGGAGAT 1018  
Qy 801 TAGACTGAAGCGATTACAATGCA 823  
Db 1019 TGGATTGATTTTACGATTGTA 1041

RESULT 9

AAC33324

ID AAC33324 standard; DNA; 1274 BP.

XX

XX

AC AAC33324;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 2614.

XX

KW Hybridisation assay; genetic mapping; gene expression control;

XX

KW protein identification; signal transduction pathway; metabolic pathway;

XX

KW promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123380P.  
PR 03-MAR-1999; 99US-0123548P.  
PR 25-MAR-1999; 99US-0125788P.  
PR 29-MAR-1999; 99US-0126264P.  
PR 01-APR-1999; 99US-0126785P.  
PR 06-APR-1999; 99US-0127462P.  
PR 08-APR-1999; 99US-0128214P.  
PR 16-APR-1999; 99US-0128714P.  
PR 19-APR-1999; 99US-0129845P.  
PR 21-APR-1999; 99US-0130077P.  
PR 23-APR-1999; 99US-0130449P.  
PR 28-APR-1999; 99US-0130510P.  
PR 30-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.

PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	08-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	09-OCT-1999;	99US-0158029P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159295P.
PR	20-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159329P.
PR	20-JUL-1999;	99US-0144336P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144337P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144338P.	PR	14-OCT-1999;	99US-0159637P.
PR	21-JUL-1999;	99US-0144814P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0145086P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160741P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160814P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160981P.
PR	26-JUL-1999;	99US-0145226P.	PR	22-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145276P.	PR	23-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	23-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0145951P.	PR	25-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	28-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149369P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149923P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151068P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
<p>Query Match 25.0%; Score 255; DB 3; Length 1274;            Best Local Similarity 59.6%; Pred. No. 4.2e-59;            Matches 462; Conservative 0; Mismatches 290; Indels 23; Gaps 1;</p>					
QY	68	ACGCTAGCTTCTGTGAGAGAGGATTGGTTAGGCGAAGATACCATCATTTATGTCTTC	127		
DB	327	ACTCTTGAAGGATTAGAACTCTTTGATCCGTCAAGAGGACAGCATTTATTTGGGCTA	386		
QY	128	ATTGAGAGAGCAAGTTCCTAGCAATTCACACCTATGATGAAAGTATGCTCAAATC	187		
DB	387	TTGGAGAGAGCAAGTACTGTTTACAATGCTGATCTATGATCCTACTCTTTGACATG	446		
QY	188	CAGGGTTTTTGTGGCTCATTTGGTGAATTTGTTTAAAGATACAGAGCCATTCAAGCT	247		
DB	447	GATGGTTTCAATGGTCTTTGGTTGAGTACATGGTTAAGGCATGAGAGCTTCAAGCT	506		
QY	248	AAGGCTGGAAGATACAAACCCCTGAAGAAAGCGCTTCTTCCAGAAATTTACACCA	307		
DB	507	AAGGTTGGTAGGTTAAGAGTCTCTGATGAACATCTCTTCTTCCCTGATGATCTACAGAG	566		
QY	308	TCAATTGTGCCATCTTACTCTTTCAACAGTTTTTGGCATCTGGTGCTCTCAATTAC	367		
DB	567	CCTATGTTGCCCTCTCTTCAGTACCACAAAGGTGTGCAATTTTGTCTGATTCGATAAC	626		
QY	368	ATAACAAGTCCATCTGGAAATGTAATTTCAAGAGTCTTCTTCCATTTGCTTACTTCG	427		
DB	627	ATAACAAGATATGGAACATGTACTTCAGAGACCTTGTTCAGAGACTTGTGAAGAA	686		
QY	428	GGTGATGAGCACTATGCGCAAACTGCGCAATGACCTTCAATTTATTTGAGTCCATC	487		
DB	687	GGGATGATGGTAACCTACGGCTCAACAGCTGCTGTGAGCTATCTGCTTCAGTGTCTC	746		
QY	488	TCTAGAAGGATTCACCTATGGAAGTTTGTAGTGAGTGAATTTGATGAGTGAATTCAGGAGTCTCCTCAA	547		



ABA91399/c

ID ABA91399 standard; DNA; 1207 BP.

XX AC ABA91399;

XX DT 08-APR-2002 (first entry)

XX DE Arabidopsis chorismate mutase-1 antisense polynucleotide.

XX KW Chorismate mutase-1; enzyme; plant; antisense; ss.

XX XX Arabidopsis thaliana.

XX OS WO200202798-A2.

XX PN 10-JAN-2002.

XX XX 22-JUN-2001; 2001WO-US020104.

XX XX 05-JUL-2000; 2000US-00610040.

XX XX (PARA-) PARADIGM GENETICS INC.

XX PI Boyes DC, Davis KR, Woessner JP, Goriach J, Hamilton CM;

XX PI Hofferman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;

XX PI KJentrup S;

XX DR WPI; 2002-154754/20.

XX PT Identifying compounds useful for modulating and inhibiting plant growth

XX PT and development, by using chorismate synthase and chorismate mutase as

XX PT targets which are essential for plant growth.

XX PS Claim 35; Page 71; 73pp; English.

XX The present sequence is that of an antisense oligonucleotide that is  
 CC specific for Arabidopsis thaliana chorismate-mutase-1 (CM) mRNA. CM  
 CC catalyzes the conversion of chorismate to prephenate in a branch of the  
 CC shikimate pathway leading to phenylalanine and tyrosine biosynthesis. It  
 CC is essential for plant growth; inhibition of CM gene expression in plant  
 CC seedlings results in severe chlorosis, reduced growth and developmental  
 CC abnormalities. The invention provides compositions for the modulation of  
 CC plant growth or development using CM antisense and sense polynucleotides,  
 CC double-stranded RNA and ribozymes, and related expression cassettes and  
 CC vectors. The compositions are particularly useful for the modulation and  
 CC inhibition of plant growth, and for generating male sterile plants or  
 CC plants that produce seedless fruit

XX SQ Sequence 1207 BP; 354 A; 278 C; 232 G; 343 T; 0 U; 0 Other;

XX Query Match 25.0%; Score 254.8; DB 6; Length 1207;

XX Best Local Similarity 57.4%; Pred. No. 4.6e-59;

XX Matches 494; Conservative 0; Mismatches 342; Indels 24; Gaps 1;

QY 68 ACGGTAGCTTCTGTGAGAGAGATTGGTTAGGCAAGAGGATACCATCATTTATGGTCTC 127

DB 955 ACTCTGAAGGTATTAGAACTCTTTCATCGCTCAAGAGGACAGCATTTATTTGGGCTA 896

QY 128 ATTGAGAGAGCCAGTTCCTAGCAATCTCAGACCTATGATGAAAAGTATGCTCAATC 187

DB 895 TTGAGAGAGAGCCAGTACTGTGTAATGCTGATATCTATGATCCTCTGCTTTTGACATG 836

QY 188 CAGGTTTTTGTGGCTCATTTGGTGGAAATTTTGTGTAAGAAATACAGAGGCCATTCAGCT 247

DB 835 GATGTTTCAATGGTCTTTGTTGAGTACATGTTAAAGGCACTGAGAAGCTTCAGCT 776

QY 248 AAGCTGGAAGATACAAAACCCCTGAGAAAACCGCTTCTTCCAGAAAATTTACACCA 307

DB 775 AAGTTGGTAGGTTAAAGATCCTGATGAACATCCTTTCTCCCTGATGATCTACCAGAG 716

QY 308 TCAATTGGCCATCTTACTCTCTCAACAGATTTTGGATCTCTGCTGCTGCTCAATTAAC 367

DB 715 CCTATGTCCTCTCTCTCAGTACCCAAAGGTGTGCAATTTGCTGCTGATTCGATNAAC 656

QY 368 ATAAACAAGTCCATCTCTGGAATATGTTATTCAGAGTTACTTCATTGCTTCTACTTCG 427  
 DB 655 ATAAACAAGAAGATATGGAACATGTACTTCAGAGACCTTGTTCAGAGATTGTGAAGAA 596  
 QY 428 GGTGATGATGGCAACTATGCGCAAACTGCAGCTTAATGACCTTTTCATTTATTCAGTCCATC 487  
 DB 595 GGCATGATGGTAACTACCGCTCAACAGCTGCTGTGACGCTATCTGCTTTCAGTGTCTC 536  
 QY 488 TCTAGAGAGATTCACTATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCTCAA 547  
 DB 535 TCAAAGAGAAATCCATTACCGTAAATTTGTGAGAGCTTAAATTTCAAGCCCTCACCCGAA 476  
 QY 548 GACTACGAGCCTTTAATTCGAGCTAAGGATAAAGAGGATTCATGAAAATTTGTGACATTT 607  
 DB 475 GCATACGAGTCCGCCATCAAGCACAAGATAAGGATCGACTGATGATATGCTGACATTC 416  
 QY 608 ACAAGCGTTGAAGAGACGGTGAAGAGAGATTGAAAAGAGCGTGTGCTGTTTGGGCGAG 667  
 DB 415 CCGACTGTGGAAGATGCGATAAAGAGAGAGATTGAGATGAAAACCCGAAATACACGGCAA 356  
 QY 668 GAAGTGAATCTTTAAACAGTGTGATGACATGACATGAAAA----- 705  
 DB 355 GAAGTGAAGATTGGGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 296  
 QY 706 --CCGTAATTTGATCCATCAGTGGCTTCTAGCTTGTACAAAATTTGGGTGATACCTCTC 763  
 DB 295 GTTTACAAAATCAGTCCGATCTTAGTTGGTGACTTATATGGAGATTGGATCATGCTTTTA 236  
 QY 764 ACCAAGGAGGTTTCAGGTTGAGTACCTCTTGGCCGCTCTAGACTGAGGCAATTACATGCA 823  
 DB 235 ACAAAGAGGTTCAAGTGGGACTTGTCTCAGAACTGGAGCTGAGGCAACCAACAAATA 176  
 QY 824 GTTAGAATTTAGAAAGATGGAAGATGAATATGATGTTGTTTCAATGATTAAGCTCTT 883  
 DB 175 AACAAATATGCTTTGCTAGTACAGTAGAAGGCTTTTGAATGCTCTTTGGTTTTTTTTT 116  
 QY 884 AAGTGATCTTTTATGCCAA 903  
 DB 115 TTACTTTACAATATTTCTAA 96

RESULT 12

ABZ12932

ID ABZ12932 standard; DNA; 1005 BP.

XX AC ABZ12932;

XX DT 21-JAN-2003 (first entry)

XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 737.

XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX OS Arabidopsis thaliana.

XX PN WO200216655-A2.

XX PD 28-FEB-2002.

XX PF 24-AUG-2001; 2001WO-US026685.

XX PR 24-AUG-2000; 2000US-0227866P.

XX PR 26-JAN-2001; 2001US-0264647P.

XX PR 22-JUN-2001; 2001US-0300111P.

XX PA (SCRI) SCRIPPS RES INST.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Harper JF, Kreps J, Wang X, Zhu T;

XX DR WPI; 2002-304127/34.

Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.

Claim 144; SEQ ID NO 737; 577pp + Sequence Listing; English.

The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an *Arabidopsis thaliana* stress regulated gene (AB212196-AB217574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

Sequence 1005 BP; 287 A; 205 C; 235 G; 278 T; 0 U; 0 Other;

Query Match 24.6%; Score 251; DB 6; Length 1005;  
Best Local Similarity 59.6%; Pred. No. 4.7e-58;  
Matches 456; Conservative 0; Mismatches 285; Indels 24; Gaps 1

QY 68 ACGCTAGCTTCTGTGAGAGAGATTGGTTAGGCAAGAGGATACATCATTTATGTCTC 127  
DB 241 ACTCTGAAGGTATTAGAACTCTTTGATCCGTCAAGAGGACGACATTATTTGGCTA 300

QY 128 ATTGAGAGACCAAGTTCCCTAGCAATTTCTCACACTATGATGAAAGTAGTGTCAATC 187  
DB 301 TTGGAGAGAGCCAAAGTACTGTGTACAACTGCTGATCTATTGATCCTACTGCTTTGACATG 360

QY 188 CAGGGTTTTTGGGCTTCATTGCTGGGAAATTTGTTTGAAGAATACAGAGGCCATTTCAAGCT 247  
DB 361 GATGGTTTCAATGGTTCTTTGGTTGGGTACATGGTTTAAAGCACTGAGAGCTTTCAGCT 420

QY 248 AAGGCTGGAAGATACAAAAACCTCGAAGAAACCCCTTCTTCCAGAAATTTACCACCA 307  
DB 421 AAGGTTGGTAGGTTTAAAGAGTCTTGATGAAACATCTTCTCCGTGATGATCTACCAGAG 480

QY 308 TCBAATTTGCCATCTTACTCCCTTCAACAGATTTTGGATCCTGGTGCCTGCTTCAATTAAC 367  
DB 481 CCTATGTTGCCCTCCCTCTCAGTACCACCAAGGTGTTCGATTTTGTCTGATGATCGATTAAC 540

QY 368 ATAAAACAAGTCCATCTGGAATAATGTATTTCAAAGAGTTACTTCCATTTGCTGTACTTCG 427  
DB 541 ATAAAACAAGATATGGAACATGTACTTCAGAGACCTTGTTCACAGACTTGTGAAGAAA 600

QY 428 GGTGATGATGGCAACTATGCGCAAACTCGACGTAAATGACCTTTCATTTATTTGAGTCCATC 487  
DB 601 GCGCATGTATGGTAACTACGGCTCAACAGCTCTCTGTGACGCTATCTGCCCTTCAGTGTCTC 660

QY 488 TCTPAGAAGGATTCACATCTATGGAAGTTTCTAGCTGAGGTGAAATTCAGGGATGCTCTCTCAA 547  
DB 661 TCMAAGAGAAATCCATTTACGGTAAATTTGTTGCAAGACTAATTTCAAGCCTCACCCGAA 720

QY 548 GACTACGAGCCCTTTAATTCGAGCTTAAGGATAAGAGAGATTGATGAAATTTGTTGACATTT 607  
DB 721 GCATACGAGTCCGCCATCAAGACCAAGATTAAGGATGCACTGTATGGATATGCTGACATTC 780

QY 608 ACAAGCGTTTGAAGAGACGCTGAGGAAGAGAGTTGAAAGAAAGCGGTGTGGTGTTTGGGCA 667  
DB 781 CCGACTGTGGAGATGCGATTAAGAAAGAGAGTTTGATGAAACCCGAAACATACGGGGCAA 840

QY 668 GAAGTGAATCTTAAACAGTGAATGACAAATGACAAATGAAAA ----- 705  
DB 841 GAAGTCAAAGTTGGGATGGAGGAGAAAGAAAGAAAGAAAGAAAGGGAATGAATCTCAT 900

QY 706 --CCGTAAATTTGATCCATCAGTGGCTTCTAGCTTGATCAAAATTCGGGTGATACCTCTC 763  
DB 901 GTTTACAAAATCAGTCCGATCTTAGTTGGTGACTTATATGGAGATGATGATCATGCCCTTTA 960

QY 764 ACCAAGAGGTTTCAGGTTGAGTACCTCTTTGGCCGCTCTAGACTGA 808

Sequence 1223 BP; 341 A; 284 C; 312 G; 286 T; 0 U; 0 Other;  
Query Match 23.5%; Score 239.4; DB 7; Length 1223;  
Best Local Similarity 56.7%; Pred. No. 7.5e-55;  
Matches 463; Conservative 0; Mismatches 351; Indels 3; Gaps 1;  
QY 2 CACGAGAGCTTAATTTCAAGATGCCAAAGCAGACAGCAACAAAGTCTCTGATCTTCGGGAAT 61  
DB 189 CGCGCGCCACAACTCCGCAACGCCATAGCTAAGGTAGAGAGGTTGATCGAAGTGC 248  
QY 62 GTGTACAGCTAGCTTCTGTGAGAGAGATTTGGTGGCAAGAGATACATCAATTTAT 121  
DB 249 ATATGACATTTGGATAGCATCAGACAGCTTTGATTTAGACTAGAGAGCAGCATCATATTT 308  
QY 122 GGTCTCATTTAGAGAGCCCAAGTTCCTAGCAATCTCACACCTATGATGAAGTATGCT 181  
DB 309 GGCCTTTTGGAGAGACACAGTTTGTACATGCTGATACATATGATAGCAATGCTTTC 368  
QY 182 CAATCCAGGTTTTTGGTGTCTATTTGGTGGAAATTTGTTTGAAGATACAGAGGCCATT 241  
DB 369 CACATGGATGTTTGGAGGATCTTTGTTGTAATATATAGTAGAGAACTGAAAGCTC 428  
QY 242 CAAGCTAGGCTGGAGATACAAACCTCAAGAAACGCTTCTCCAGAAAATTTA 301  
DB 429 CATGACAGGTTGGAGATACAGAGCCAGCATGAGCACCTTTCTTTTCCAGGATCTG 488  
QY 302 CCACCATCAATTTGCTCATCTTACTCTTCAACAGTTTTTGCATCTCTGCTGCTTCA 361  
DB 489 CCTGAGCCCGTTGGCCACCTATCAATACCAAGGTTTTCATCCATTCGTGATCT 548  
QY 362 ATTAACATAACAGTCCATCTGAAATGATTTTCAAGATTTACTTCAATGCTTGCT 421  
DB 549 ATCAATATCAACAAGAGATTTGAAATGATTTTGTATGAACTTCTTCCAAGATTGGTG 608  
QY 422 ACTTCGGGTGATGATGCAACTATGCGCAACTGATCAATGACCTTTCATTTATTGCGAG 481  
DB 609 AAGAAGGAAGTGAATGATGCTGGATCCAGTCTCTTTGTGCACAACTGCTGTCAG 668  
QY 482 TCCATCTCTAAGAGATTCATATGGAAGTTTGTAGCTAGGTGAAATTCAGGATGCT 541  
DB 669 GCACTCTCCAGAGGATCCACTATGGAAGTTTGTGCGAGAGGCCAAAGCCAGATCTC 728  
QY 542 CCTCAAGACTACGAGCTTTTAATTCGAGCTAAGATTAAGAAGATTGATGAATTTTG 601  
DB 729 CTTGAGCTTACGCGCAGCCATTAATGCCCAAGCCGATCACTCACTGATCACTTCTC 788  
QY 602 ACATTTCAAGCGTTGAAGAGACGCTGAGGAAGAGATTGAAAGAAGGCTGTGTTGTT 661  
DB 789 ACATATGAACGGTGGAGCGTGCTATCGAACACAGGTTGGAGGCCAAAGCCAGATCTC 848  
QY 662 GCGCAGGAGTGAATCTTAACAGTATGATGACATGACATGAA---AACCGTAAATTTGAT 718  
DB 849 GGGCAAGAGGTGAACATTTGGTGTGAAGGACAAACGCGACCCACCGGTCTACAAATCAGG 908  
QY 719 CCATCAGTGGCTTCTAGCTGTACAAAATTTGGGTGATACCTCTCACCAAGGAGTTTCAG 778  
DB 909 CCGAGCTGTGTCGCGAGCTGTACAGCTACAGATCATGCGCTAACCAAGGAGTTGAG 968  
QY 779 GTTGAGTACCTCTTGGCGCGCTCTAGACTGAAGGCAAT 815  
DB 969 CTCGCGTACTTGTGTTAAGAGGCTGGATTTGATGTTT 1005

## RESULT 14

ABX93059 standard; cDNA; 780 BP.  
ID ABX93059  
XX  
AC  
XX  
XX  
DT 22-MAY-2003 (first entry)  
XX  
DE Wheat chorismate mutase cDNA.  
XX

Query Match 22.5%; Score 229.4; DB 7; Length 780;  
Best Local Similarity 63.6%; Pred. No. 3.3e-52;  
Matches 391; Conservative 0; Mismatches 206; Indels 18; Gaps 2;  
QY 212 GAATTTGTTGTTAAGATACAGAGCCCATTTCAAGTAAAGGCTTGAAGATACAAAACCT 271  
DB 10 GAGTTCTTCGTTCCGGAAGCCGAGGCTCTGACGCAAGGCTGGACACTATCAAAGGCA 69  
QY 272 GAAGAAACGCTCTTCCAGAAATTTACACCATCAATTTGCCATCTTACTCTTC 331  
DB 70 GAAGATGTTCCATTTCTTCTCAAGATCTTCCACCCTGCTTTTCTACCAAGGTCG 129

Wheat; plant; gene; ss; chorismate mutase; aromatic amino acid; diet;  
biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;  
phenylalanine; prephenate dehydrogenase; herbicide; food crop.  
Triticum aestivum.

Key Location/Qualifiers  
CDS 1..624  
/tag= a  
/product= "Chorismate mutase"  
/partial  
/note= "No start codon shown"

US2002184658-A1.

05-DEC-2002.

03-DEC-1999; 99US-00454279.

04-DEC-1998; 98US-0110845P.

(FALC/) FALCO S C.  
(FAMO/) FAMODU O O.  
(LEEF/) LEE J.

Falco SC, Famodu OO, Lee J;

WPI; 2003-328651/31.

P-PSDB; ABU08095.

New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
nucleic acids, useful for facilitating design and/or identifying  
inhibitors of those enzymes that may be used as herbicides and for  
producing antibodies.

Claim 3; Page 23; 32pp; English.

The invention discloses isolated polynucleotides encoding chorismate  
mutase polypeptides. Aromatic amino acids must be included in the diets  
of animals. In the aromatic amino acid biosynthetic pathway chorismate is  
converted to anthranilate during tryptophan biosynthesis and is converted  
to prephenate, the branch point for tyrosine and phenylalanine  
biosynthesis. Chorismate mutase catalyzes the conversion of chorismate to  
prephenate. Also disclosed are methods for selecting an isolated  
polynucleotide that affects the level of expression of a tyrosine  
biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
evaluating compounds for their ability to inhibit the activity of a  
tyrosine biosynthetic enzyme. The polypeptides can be used produce  
antibodies. Chorismate mutase and prephenate dehydrogenase are good  
targets for herbicides that will not affect animals, and overexpression  
of these enzymes may be used to increase the content of aromatic amino  
acid in food crops. The polypeptides may also be used to design and/or  
identify inhibitors of those enzymes that may be used as herbicides. The  
nucleic acids may be used to create transgenic plants, as probes for the  
genetic and physical mapping of the genes and as markers for traits  
linked to those genes. The sequence presented is a wheat chorismate  
mutase cDNA

Sequence 780 BP; 256 A; 146 C; 169 G; 209 T; 0 U; 0 Other;





Db 1083 GACTGGATTAAATG 1095

Search completed: September 24, 2004, 23:23:37  
Job time : 482 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 23:13:06 ; Search time 97 Seconds  
(without alignments)  
5835.566 Million cell updates/sec

Title: US-10-624-061-15  
Perfect score: 1020  
Sequence: 1 gcacgagagcttaatttca.....atcttcaaaaaaaaaaaaaa 1020

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1020	100.0	1020	4	US-09-454-279-15
2	465.6	45.6	525	4	US-09-454-279-5
3	327.2	32.1	1006	4	US-09-610-040-3
4	327.2	32.1	1006	4	US-09-610-040-9
5	257.4	25.2	1231	4	US-09-454-279-13
6	254.8	25.0	1207	4	US-09-610-040-2
7	254.8	25.0	1207	4	US-09-610-040-8
8	239.4	23.5	1223	4	US-09-454-279-11
9	229.4	22.5	780	4	US-09-454-279-17
10	211.4	20.7	1217	4	US-09-610-040-4
11	211.4	20.7	1217	4	US-09-610-040-10
12	206.2	20.2	579	4	US-09-454-279-9
13	98.8	9.7	5176	4	US-09-610-040-6
14	73.6	7.2	5099	4	US-09-610-040-5
15	69.2	6.8	552	4	US-09-454-279-1
16	58.4	5.7	601	4	US-09-454-279-3
17	51.8	5.1	7218	1	US-08-232-463-14
18	38	3.7	3001	4	US-09-539-333D-151
19	38	3.7	3001	4	US-09-539-333D-155
20	37.8	3.7	2013	4	US-09-596-196-3
21	37.5	3.7	8093	4	US-10-204-708-32
22	36.2	3.5	7664	4	US-10-204-708-84
23	36.2	3.5	580073	4	US-08-545-528D-1
24	36	3.5	1894	3	US-09-004-731-29
25	36	3.5	1894	3	US-09-004-731-31
26	36	3.5	1894	3	US-09-032-215-3
27	36	3.5	1894	3	US-08-749-699-29

C 28 36 3.5 1894 3 US-08-749-699-31 Sequence 31, Appl  
29 36 3.5 1894 4 US-09-004-729-29 Sequence 29, Appl  
30 36 3.5 1894 4 US-09-004-729-31 Sequence 31, Appl  
31 36 3.5 10619 4 US-10-204-708-3 Sequence 3, Appl  
32 35.8 3.5 640681 4 US-09-790-988-1 Sequence 1, Appl  
33 35.6 3.5 903 4 US-09-533-029-13 Sequence 13, Appl  
34 35.4 3.5 1602 4 US-09-672-494-1 Sequence 1, Appl  
35 35.4 3.5 1718 4 US-09-215-450-17 Sequence 17, Appl  
36 35.2 3.5 640681 4 US-09-790-988-1 Sequence 1, Appl  
37 35 3.4 349 4 US-08-961-527-335 Sequence 335, App  
38 35 3.4 903 4 US-09-489-039A-581 Sequence 581, App  
39 35 3.4 2700 3 US-09-022-875-1 Sequence 1, Appl  
40 35 3.4 58407 4 US-08-916-421B-2 Sequence 2, Appl  
41 35 3.4 1830121 4 US-09-557-884-1 Sequence 1, Appl  
42 35 3.4 1830121 4 US-09-643-390A-1 Sequence 1, Appl  
43 34.8 3.4 822 4 US-09-543-681A-3973 Sequence 3973, Ap  
44 34.8 3.4 6317 4 US-10-204-708-11 Sequence 11, Appl  
45 34.4 3.4 1290 4 US-09-543-681A-3371 Sequence 3371, Ap

ALIGNMENTS

RESULT 1  
US-09-454-279-15  
; Sequence 15, Application US/09454279  
; Patent No. 662798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB4299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 15  
; LENGTH: 1020  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-454-279-15

Query Match 100.0%; Score 1020; DB 4; Length 1020;  
Best Local Similarity 100.0%; Pred. No. 1.2e-261;  
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCACGAGAGCTTAATTTCAGAAATGCCCAACGACGAGNACAAAGTCTCTGATTCGGGA 60  
DB 1 GCACGAGAGCTTAATTTCAGAAATGCCCAACGACGAGNACAAAGTCTCTGATTCGGGA 60  
QY 61 TGTGTACACGCTAGCTTCTGTGAGAGAGATTGGTTAGCAAGAGGATACCATCATTTA 120  
DB 61 TGTGTACACGCTAGCTTCTGTGAGAGAGATTGGTTAGCAAGAGGATACCATCATTTA 120  
QY 121 TGGTCTCAATGAGAGAGCAAGTCCCTAGCAAAATCTCAACCTATGATGAAAAGTATGC 180  
DB 121 TGGTCTCAATGAGAGAGCAAGTCCCTAGCAAAATCTCAACCTATGATGAAAAGTATGC 180  
QY 181 TCAATCCAGGCTTTTGTGGTCTCAATGGTGAATTTGTGTAGAAATACAGAGGCCAT 240  
DB 181 TCAATCCAGGCTTTTGTGGTCTCAATGGTGAATTTGTGTAGAAATACAGAGGCCAT 240  
QY 241 TCAAGCTAAGGCTTGAAGATACAAAACCCCTGAAGAAAACGGCTTCTCCAGAAAATTT 300  
DB 241 TCAAGCTAAGGCTTGAAGATACAAAACCCCTGAAGAAAACGGCTTCTCCAGAAAATTT 300  
QY 301 ACCACATCAATTTGCCATCTTACTCTCTTCAACAGATTTTTCATCTCGTGTCTCTTC 360  
DB 301 ACCACATCAATTTGCCATCTTACTCTCTTCAACAGATTTTTCATCTCGTGTCTCTTC 360

361 AATTAAACATAAACAAGTCCATCTGGAAATATTTTCAAGAGTTACTTCCATTGCTTGC 420  
361 AATTAAACATAAACAAGTCCATCTGGAAATATTTTCAAGAGTTACTTCCATTGCTTGC 420  
421 TACTTCGGGTGATGATGGCAACTATGCGCAAACTGAGCTTAATGAGCTTTTCAATTTGCA 480  
421 TACTTCGGGTGATGATGGCAACTATGCGCAAACTGAGCTTAATGAGCTTTTCAATTTGCA 480  
481 GTCCATCTCTAGAAGGATTCACATATGGAAGATTTTGTAGCTGAGGTCAAAATTCAGGATGC 540  
481 GTCCATCTCTAGAAGGATTCACATATGGAAGATTTTGTAGCTGAGGTCAAAATTCAGGATGC 540  
541 TCCTCAAGACTAGAGCTTTAATTCGAGCTAAGGATAAGAGGATTTGATGAATTTGTT 600  
541 TCCTCAAGACTAGAGCTTTAATTCGAGCTAAGGATAAGAGGATTTGATGAATTTGTT 600  
601 GACATTTACAAGCGTTGAAGAGACGGTGAAGAGAGAGTTGMAAAGAGCGTGTGTGTT 660  
601 GACATTTACAAGCGTTGAAGAGACGGTGAAGAGAGAGTTGMAAAGAGCGTGTGTGTT 660  
661 TGGCAGGAGAGTGAATCTTAACAGTGAATGAATGAATGAATGAATGAATGAATGAATGAAT 720  
661 TGGCAGGAGAGTGAATCTTAACAGTGAATGAATGAATGAATGAATGAATGAATGAATGAAT 720  
721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGT 780  
721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGT 780  
781 TGAGTACCTCTGGCGGCTAGACTGAAGGCAATTAATGAGCTTTAGAGTTTGAAGAA 840  
781 TGAGTACCTCTGGCGGCTAGACTGAAGGCAATTAATGAGCTTTAGAGTTTGAAGAA 840  
841 TGAAGATGAATATGATGTTGTTGTTCAATGATTAAGCTCTTAAGTGAATCTTTATTC 900  
841 TGAAGATGAATATGATGTTGTTGTTCAATGATTAAGCTCTTAAGTGAATCTTTATTC 900  
901 CAACCTCATGCTGTTGATTCAGAAATATTTTGTAGCTATAGATTTTACCTTTAT 960  
901 CAACCTCATGCTGTTGATTCAGAAATATTTTGTAGCTATAGATTTTACCTTTAT 960  
961 TTTTCATCGCTTTATTAGAAAGGCAATTTATATCATGATCTTCAAAAAA 1020  
961 TTTTCATCGCTTTATTAGAAAGGCAATTTATATCATGATCTTCAAAAAA 1020

## RESULT 2

US-09-454-279-5  
; Sequence 5, Application US/09454279  
; Patent No. 6627798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 5  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (432)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (514)  
; FEATURE:  
; NAME/KEY: unsure

; LOCATION: (516)  
US-09-454-279-5

Query Match Similarity 45.6%; Score 465.6; DB 4; Length 525;  
Best Local Similarity 95.4%; Pred. No. 1.9e-114;  
Matches 499; Conservative 0; Mismatches 22; Indels 2; Gaps 2;

QY 8 AGCTTAATTTCAAGATGGCCAAAGCAGCAGAAACAAAGTCTCTGATTTGGGAATGTGTAC 67  
DB 1 AGCTTAATTTCAAGATGGCCAAAGCAGCAGAAACAAAGTCTCTGATTTGGGAATGTGTAC 60  
QY 68 ACGTAGCTTCTGTGAGAGAGGATTTGGTTAGGAAGAGGATACCATTTATGGTCTC 127  
DB 61 ACGTAGCTTCTGTGAGAGAGGATTTGGTTAGGAAGAGGATACCATTTATGGTCTC 120  
QY 128 ATTGAGAGAGCCAGTTCCTTAGCAATTTCTCACACCTTATGATGAAAAGTATGCTCAAATC 187  
DB 121 ATTGAGAGAGCCAGTTCCTTAGCAATTTCTCACACCTTATGATGAAAAGTATGCTCAAATC 180  
QY 188 CAGGGTTTTGTGGCTCATTGGTGGAAATTTGTTTAAAGAAATACAGAGGCCATTCAGCT 247  
DB 181 CAGGGTTTTGTGGCTCATTGGTGGAAATTTGTTTAAAGAAATACAGAGGCCATTCAGCT 240  
QY 248 AAGCTCGAAGATACAAAAACCTTGAAGAAACGCTTCTTCCAGAAAATTTTACCACCA 307  
DB 241 AAGCTCGAAGATACAAAAACCTTGAAGAAACGCTTCTTCCAGAAAATTTTACCACCA 300  
QY 308 TCAATTTGCAATCTTACTCTTCAAAACAGTTTTTGTGATCTGGTCTGCTTCAATTAAC 367  
DB 301 TCAATTTGCAATCTTACTCTTCAAAACAGTTTTTGTGATCTGGTCTGCTTCAATTAAC 360  
QY 368 ATAACAGTCCATCTGGAATATGTTTCAAGAGCTTACTTCCATTTGCTGCTTCTG 427  
DB 361 ATAACAGTCCATCTGGAATATGTTTCAAGAGCTTACTTCCATTTGCTGCTTCTG 420  
QY 428 GGTGATGATGGCAACTATGCGCAAACTGCGAGCTTAATGACCTTTTCAATTTGCACTC 487  
DB 421 GGTGATGATGGCAACTATGCGCAAACTGCGAGCTTAATGACCTTTTCAATTTGCACTC 479  
QY 488 TCTAGAGGATTCACATATGGAAGATTTGTAGCTGAGGTGAAT 530  
DB 480 TCTAGAGGATTCACATATGGAAGATTTGTAGCTGANGNGAAAT 521

## RESULT 3

US-09-610-040-3  
; Sequence 3, Application US/09610040  
; Patent No. 6465217  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Gorlach, John  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Klotz, Andreas  
; APPLICANT: Zayed, Adel  
; APPLICANT: Ascenzi, Robert  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE  
; FILE REFERENCE: 9128.14  
; CURRENT APPLICATION NUMBER: US/09/610,040  
; CURRENT FILING DATE: 2000-07-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 1006  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana

Query Match 32.1%; Score 327.2; DB 4; Length 1006;  
Best Local Similarity 64.2%; Pred. No. 1.5e-77;

Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
QY 22 AATGGCCAAAGCAGAGCAAAAGTCTGATCTCGGAATGTGTACACGCTAGCTTCTGT 81  
DB 69 AAGAGTCTTCCGAATCGGATTCGGGTTCTGGTTCTTCCAAATGTACTGAGTCTTGAATTAAT 128  
QY 82 GAGAGGATTTGGTTAGGCAAGAGATACCATTTATGGTCTCATTCAGAGAGCCAA 141  
DB 129 CAGAGATCGTTGANTAGGCAAGAGACACATCGTTCAGCTTGATCGAGAGACTAA 188  
QY 142 GTTCCCTAGCAATCTCACACCTATGATGAAGAAAGTATGCTCAAAATCCAGGTTTGTGG 201  
DB 189 GTTCCCACTCAATCTCCTGTTTCGAGGAATCTCGTTGTCTAGATTTCTGAAATTTCTC 248  
QY 202 CTCAATGGTGAATTTGTTTGAATATACAGAGCCATTCAGCTTAAGGCTCGAAGATA 261  
DB 249 TTCTCTCACTGATTTTCTGTCAGAGACAGAAATCATCCAGCTTAAGTAGAAGATA 308  
QY 262 CAAAAACCTGAAGAAACCCCTTTCTCCAGAAATTTTACCAACCATCAATTTGCCATC 321  
DB 309 TGAATACCCGAGAGAAATCCTTTCTTCTTGAACATTTCTCACTCGGTTTTTCTAC 368  
QY 322 TTACTCCTTCAACAGATTTTGCATCCTGGTGTCTTCAATTAACATTAACAGTCCAT 381  
DB 369 GCACAAATATCCATCGGCTTTGCACCCCTAAGGCTCTATCTGTAAACATTAACAAACAAT 428  
QY 382 CTGGAATATGTAATTCAGAGATTTACTTCCATTGCTTCTTCTGCTGATGATGCA 441  
DB 429 CTGGATATTTACTTTAAGAAATTTGCTTCTTCTTGAACATTTCTTCTGAGAGATTTCA 488  
QY 442 CTATGCGCAACTCGCAGTAAATGACCTTTTCAATTTGCAAGTCCATCTCTAGAGAGATTCA 501  
DB 489 CTATCCATCAACTCGTCTAGTATCTCGCTGTTTCAAGCTCTTTTCGAGAGAGATTCA 548  
QY 502 CTATGGAAAGTTTCTAGCTAGGTTGAATTCAGGATGCTTCTCAAGACTACGAGCCTTT 561  
DB 549 CTACGGTAATTTGAGTCTAGGTTCAATTCAGAGATGCTTCCACAGATTTACGAGCTGC 608  
QY 562 AATTCGAGCTAAGGATAAGAGAGATTCATGAATTTGTTGACATTTACAGCGTTGAAGA 621  
DB 609 GATTTCGGCTCAGGATAGAGAGCTTTGATGAAGCTGTGACGTTTGAGAAAGTAGAAGA 668  
QY 622 GACGCTGAGGAGAGACTTGAAGAGAGGCTGTGCTGTTTGGCAGGAGAGTGAATCTTAA 681  
DB 669 AATGGTTAAGAGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 728  
QY 682 -----CAGTGATGACATGCAATGAAGAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735  
DB 729 CTCTGGCTATGCGATGAGAGTGAAGAGAGATATAAAGTGGATCCATTGCTTGGCTCTCG 788  
QY 736 CTGTTACAAAATTTGGGTGATACCTCTCAACAGAGGTTTCAAGTTGAGTACCTCTTGG 795  
DB 789 CATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTTCTACG 848  
QY 796 CCGTCTAGACTGAA 809  
DB 849 TCGTCTCGATTGAA 862

RESULT 4

US-09-610-040-9/c

Sequence 9, Application US/09610040

Patent No. 6,465,217

GENERAL INFORMATION:

- APPLICANT: Boyes, Douglas
- APPLICANT: Davis, Keith
- APPLICANT: Woessner, Jeffrey
- APPLICANT: Goriach, John
- APPLICANT: Hamilton, Carol
- APPLICANT: Hoffman, Neil
- APPLICANT: Klotz, Andreas
- APPLICANT: Zayed, Adel
- APPLICANT: Ascenzi, Robert

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHESIS  
FILE REFERENCE: 9128.14  
CURRENT APPLICATION NUMBER: US/09/610,040  
CURRENT FILING DATE: 2000-07-05  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 9  
LENGTH: 1006  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-610-040-9

Query Match 32.1%; Score 327.2; DB 4; Length 1006;

Best Local Similarity 64.2%; Pred. No. 1.5e-77;

Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 AATGGCCAAAGCAGAGCAAAAGTCTGATCTCGGAATGTGTACACGCTAGCTTCTGT 81  
DB 938 AAGAGTCTTCCGAATCGGATTCGGGTTCTGGAATGTGTACACGCTTGTACTTAAT 879  
QY 82 GAGAGAGATTTGGTTAGGCAAGAGATACCATTTATGGTCTCTCATTCAGAGAGCCAA 141  
DB 878 CAGAGATCGTTGATGAGCAAGAGACACCATCGTCTTCAAGTGTACTGATTCGAGAGACTAA 819  
QY 142 GTTCCCTAGCAATCTCACACCTATGATGAAGATGCTCAATCCAGGTTTGTGG 201  
DB 818 GTTCCCACTCAATCTCCTGCTTTCGAGAAATCTCGTTGTCTAGATTTCTGAAATTTCTC 759  
QY 202 CTCATTGGTGGAAATTTGTTTGAATACAGAGCCATTCAGCTTAAGGCTGGAAGATA 261  
DB 758 TTCTCTCACTGATTTTTCGTCAGAGACAGAAATCATCCAGCTTAAGTAGAAGATA 699  
QY 262 CAAAAACCTGAAGAAACCCCTTTCTCCAGAAATTTTACCAACCATCAATTTGCCATC 321  
DB 698 TGAATACCCGAGAGAAATCCTTTCTCTTGGAGAAATCTCTCACTCGGTTTTTCTAC 639  
QY 322 TTACTCCTTCAACAGATTTTGCATCCTGGTCTGCTTCAATTAACATTAACAGTCCAT 381  
DB 638 GCACAAATATCCATCGGCTTTTGCACCCCTAAGCTCTATCTGTTAATTAACAAACAAT 579  
QY 382 CTGGAATATGTAATTTCAAAGAGTTACTTCCATTTGCTGCTACTTCCGTTGATGAGCA 441  
DB 578 CTGGGATATTTACTTTAAGAAATTTCTTCTTGTTCGAAACCTGCGCATGATGCA 519  
QY 442 CTATGGCAATCTGACGCTAATGACCTTTCATTTGAGTCCATCTCTAGAGAGATTCA 501  
DB 518 CTATCCATCACTGCTGCTAGTATCTCGCTGTTTACAAGCTCTTTCGAGAGATTCA 459  
QY 502 CTATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCAAGACTACGAGCCTTT 561  
DB 458 CTACGGTAATTTGTAGCTGAGTCAATTCAGAGATGCTCCACAAAGATTACGAGCTGC 399  
QY 562 AATTCAGCTAAGGATAAGAGAGTTGATGAATTTGTTGACATTTACAGGTTGAAGA 621  
DB 398 GATTCGCTCAGGATAGAGAGGCTTTGATGAAGCTGTGACGTTTGAGAAAGTAGAAGA 339  
QY 622 GACGCTGAGAGAGAGATTGAAAGAGGCTGTGCTGTTTGGCAGGAGTGAATCTTAA 681  
DB 338 AATGGTTAAGAGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 279  
QY 682 -----CAGTGATGACATGCAATGAAGAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735  
DB 278 CTCTGGCTATGCGATGAGAGTGAAGAGAGATATAAAGTGGATCCATTGCTTGGCTCTCG 219  
QY 736 CTGTTACAAAATTTGGGTGATACCTCTCAACAGAGGTTTCAAGTTGAGTACCTCTTGG 795  
DB 218 CATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTTCTACG 159  
QY 796 CCGTCTAGACTGAA 809  
DB 158 TCGTCTCGATTGAA 145

## RESULT 5

US-09-454-279-13

; Sequence 13, Application US/09454279

; Patent No. 6627798

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Famodu, Omolayo O.

; APPLICANT: Lee, Jian-Ming

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; FILE REFERENCE: BB1299 US NA

; CURRENT APPLICATION NUMBER: US/09/454,279

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110,845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 13

; LENGTH: 1231

; TYPE: DNA

; ORGANISM: Oryza sativa

US-09-454-279-13

## Query Match

25.2%; Score 257.4; DB 4; Length 1231;

Best Local Similarity 58.4%; Pred. No. 5.5e-59;

Matches 469; Conservative 0; Mismatches 331; Indels 3; Gaps 1;

QY 24 TGGCCAAAGCAGCAGCAAGCTCTGTTCTGGGAATGTTACAGCGTAGCTTCTGTCA 83

DB 239 TGACCAAGAGGAGAAACAGAGATAGATCAAGTGAATATGACCTTGGCAACATTA 298

QY 84 GAGAGAGATTGTTAGGCAAGAGGATPACCATATTTATGTTCTCATTTGAGAGAGCAAGT 143

DB 299 GAACCTCTTGGTTAGGCAAGAGCAGCATCATATTCAGCGCTCTTAGAGAGACACAGT 358

QY 144 TCCTAGCAATTCACACCTATGATGAAGATGCTCAAAATCCAGGTTTTTGTGCT 203

DB 359 TTGCTACAAATGCTGATATATATGATGAAGATGCTTTCCATGTGATGATTTGATGCT 418

QY 204 CATTTGTTGAATTTGTTTGAATATACAGAGCCATTCAGCTAAGGCTGGAAGATACA 263

DB 419 CTTTGTGTAATTCATGGTTAGAGAAACGAAACATACATCAACAGGTTGGAGATACA 478

QY 264 AAAACCTGAGAAAGCGCTTTCCAGAGAAATTTACCATCAATTTGGCATCTTT 323

DB 479 AGAGCCCTGATGAGCACCCTATCTTTCCGAGGATCTGCTGAACCATCTTTGCCACCTC 538

QY 324 ACTCTTCAAAAGTTTGTGATCTCTGCTGCTGCTTCAATTAACATAAAGTCCATCT 383

DB 539 TCAGTATCCAAAGTTTGTGATCTCTGCTGATCTATTAATATCAACAGGATTT 598

QY 384 GGAATGATTTCAAGAGTTACTTCCATGCTTGTCTACTCTTGGGTGATGATGCAACT 443

DB 599 GGAATGATTTTGTGATGCTTCTTCCAGATTTAGTGAAGAGGAAAGTGTGTAAT 658

QY 444 ATGCGCAAACTGAGCTAATGACCTTTCAATTTGAGTCCATCTCTAGAGGATTTCACT 503

DB 659 ATGATCCAGTCTCTTGTGACAGATCTGTTGAGCGGCTCTCCAAAGATTTCACT 718

QY 504 ATGGAAGTTTGTAGTGAAGTGAATTCAGGATGCTCTCAAGACTACGAGCTTTAA 563

DB 719 ATGTTAAGTTTGTGAGAGGCTAAGTTTCAAGAGTCTCTCAAGCTTACATGCTGGA 778

QY 564 TTCGACTAAGATTAAGAGGTTTGAATTTGATGATTTGATGATTTGATGATTTGATGAT 623

DB 779 TAATGACAGAGCTCGGATCACTTAATGACCTCTCTCACTATGAACCGGTGGAGCTG 838

QY 624 CGGTGAGGAGAGTTGAAAAGAGGCTGTGGTGTGTTGGCAGGAAGTGAATCTTAACA 683

DB 839 CTATTGAACATAGGTTGAAGCTAAGGCTTAAGATCTTTGGACAGGAGGTGATTTAGCG 898

QY 684 GTGATGACATGACAAAT---GAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTGT 740

DB 899 CTGAAGACACGGCGCTCCACCAATGTACAAGATAGGCCAGTTGGTGGCTGAAGTGT 958

QY 741 AAAAAATTTGGTGATACCTCTCAACAGGAGGTTCAAGTTGAGTACCTCTTGGCGGTC 800

DB 959 ACAGCTACAGGATCATGCGCTAACCAAGAGGTTCAAGTAGCTACTTCTGAGGAGAT 1018

QY 801 TAGACTGAAGCATTAACATGCA 823

DB 1019 TGGATTGATTGTTTACGATTGTA 1041

## RESULT 6

US-09-610-040-2

; Sequence 2, Application US/09610040

; Patent No. 6465217

; GENERAL INFORMATION:

; APPLICANT: Boyes, Douglas

; APPLICANT: Davis, Keith

; APPLICANT: Woessner, Jeffrey

; APPLICANT: Gornach, John

; APPLICANT: Hamilton, Carol

; APPLICANT: Hoffman, Neil

; APPLICANT: Kloti, Andreas

; APPLICANT: Zayed, Adel

; APPLICANT: Ascenzi, Robert

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE

; FILE REFERENCE: 9128.14

; CURRENT APPLICATION NUMBER: US/09/610,040

; CURRENT FILING DATE: 2000-07-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 1207

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-610-040-2

## Query Match

25.0%; Score 254.8; DB 4; Length 1207;

Best Local Similarity 57.4%; Pred. No. 2.7e-58;

Matches 494; Conservative 0; Mismatches 342; Indels 24; Gaps 1;

QY 68 ACCTAGCTTCTGTGAGAGGATTTGGTTAGCAGAGGATACCATTTATGCTCTC 127

DB 253 ACTCTTGAAGGATTTAGAACTCTTTGATCCGCTCAAGAGGACAGCAATTATTTGGGCTA 312

QY 128 ATTGAGAGAGCCCAAGTTCCCTAGCAATTTCTCACCTATGATGAAAGTATGCTCAATC 187

DB 313 TTGAGAGAGCCCAAGTACTGTTACATGCTGATCTATGATCTACTGCTTTTGACATG 372

QY 188 CAGGTTTTTGTGGCTCATTTGGTGGAAATTTGTTTGAAGATACAGAGGCCATTCAAGCT 247

DB 373 GATGGTTTCAATGGTTCTTTGGTTGATGATGTTTAAAGGCACCTGAGAAGCTTCAAGCT 432

QY 248 AAGCTGGAAGATACAAAAACCCCTGAAGAAAACCGCTTCTCCAGAAAATTTACCACCA 307

DB 433 AAGTTTGTAGGTTTAAAGTCTCTGATGAACATCTTTCTTCCCTGATGATCTACAGAG 492

QY 308 TCAATTTGTCATCTTACTCTCTTCAACAGATTTTTCATCTCTGGTGTCTGCTTCAATTAAC 367

DB 493 CCTATGTGCTCTCTCTTTCAGTACCCAAAGGTTTGTTCATTTGCTGCTGATTCGATAAC 552

QY 368 ATAAACAAGTCCATCTGGAATGTTTCAAGAGTACTTCCATTTGCTTCTCTCTCTCG 427

DB 553 ATAAACAAGAGATATGGAACATGTTACTTCAAGAGCTTGTTCAGAGATTTGAGAAA 612

QY 428 GTGTGATGGCAACTATGCGCAACTGACGCTTAATGACCTTTTCAATTTATGTCAGTCCATC 487

DB 613 GCGGATGATGTAACCTACGCTCAACAGCTGCTGTGACGCTATCTGCTTCAAGTCTC 672

QY 488 TCTAGAGGATTCACATATGGAAGTTTGTAGCTGAGTGAATTTAGGATGATCTCTCTCAA 547

DB 673 TCAAGAGATTCCTTACGTTAAATTTTGTGAGAGCTTAATTTTCAAGCTTCAACCCGAA 732



Matches 463; Conservative 0; Mismatches 351; Indels 3; Gaps 1;  
QY 2 CACGAGAGCTTAATTTCAAGAAATGGCCAAAGCAGCAGACAAAGTCTCTGATTCGGAAAT 61  
Db 189 CGCGGGCCAAACACTCCGCAAGCCCATAGCTAGGTAGAGAGGGTGTATCGAAGTGAC 248  
QY 62 GTGTACAGCTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGATACCATCAATTTAT 121  
Db 249 ATATTGACATTTGGATAGCATCAGACAGATTTTGAATAGACTAGAACAGACATCATATTT 308  
QY 122 GGTCTCATTCAGAGAGCCAGTTCCTAGCAATTTCTCACCTATGATGAAAGATGCT 181  
Db 309 GGCCTTTTGGAGAGAGACAGTTTGTACAACTGCTGATACATATAGTACGATGCTTTC 368  
QY 182 CAATCCAGGGTTTGTGCTCATTTGGTGGAAATTTGTTAAGAAATACAGAGGCCATT 241  
Db 369 CACATGGATGGTTTGGAGGATCTTTGGTTGAATATATAGTTAGAGAACTGAAAGCTC 428  
QY 242 CAGCTAAGGCTGAGATACAAACCCCTGAGAAAGCCCTCTTCCAGAAATTTA 301  
Db 429 CATGCAAGTTGGGAGATACAAGGCCAGATGAGACCCCTTTCTTTTCCAGGATCTG 488  
QY 302 CCACCATCAATTTGGCCATCTTACTCCTTCAACAGATTTTTCATCTGCTGCTTCA 361  
Db 489 CTTGAGCCCGGTTGCCACCTATGCAATACCCAGAGGTTTTCATCCCATCTGCTTCT 548  
QY 362 ATTACATAAACAGTCCATCTCGAAATGTATTTCAAGAGTTACTTCCATGCTTGTCT 421  
Db 549 ATCAATATCAACAAAGAGATTTGGAAATGTATTTGATGAACCTTCTTCCAGATTTGGT 608  
QY 422 ACTTCGGTGTATGCGCACTATGCGCAACTGCGCAACTGCGCACTTTCATTTATTTGCG 481  
Db 609 AAGAGAGGAGTGAATGATGCTGATCCAGTCCAGTCTTTTGTGACACACCTGCTTGAG 668  
QY 482 TCATCTCTAGAGAGTTCACATGAGAAATTTGTAGTGTAGTGTAGTGAATTCAGGATCT 541  
Db 669 GCACCTCCAGAGGATCCACTATGCGAAGTTTGTGCGAGAGCCCAAGTTTCAAGATGCC 728  
QY 542 CTTCAAGACTACGAGCTTTTAACTTCGAGCTAGGATTAAGAGGATTTGATGAATTTGTT 601  
Db 729 CTTGAAGCTTACGCCACCATATATAGCCCAAGCCGCTGATCACTCATGACCTTCTC 788  
QY 602 ACATTTACAAGCGTTGAAGAGACGCTGAGGAGAGAGATTTGAAAGAGGCTGTGTGTTT 661  
Db 789 ACATATGAACCGTGGAGCGTGTCTACGAACACAGCGGTGGAGCCCAAGATCTTC 848  
QY 662 GGCAGAGAGTGAATTTAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 718  
Db 849 GGCAGAGAGTGAACATTTGGTCTAGGACACAGCGCAGCCACCGGTCTACAAATCAGG 908  
QY 719 CCATCAGTGGCTTTAGCTTTGTACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAG 778  
Db 909 CCGAGCTTGTGCGCGAGCTGTACAGCTACAGATCATCGCTAACCCAGAGGTTTTCAG 968  
QY 779 GTTGAAGTCTTTGGCGCGCTTACACTGAAGGCATT 815  
Db 969 GTGCGGTACTTCTTAAGAGGCTGATTTGATGTTGTT 1005

## RESULT 9

US-09-454-279-17

; Sequence 17, Application US/09454279

; Patent No. 6627798

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Famodu, Omolayo O.

; APPLICANT: Lee, Jian-Ming

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; FILE REFERENCE: BB1299 US NA

; CURRENT APPLICATION NUMBER: US/09/454, 279

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110, 845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 17  
; LENGTH: 780  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
US-09-454-279-17

Query Match 22.5%; Score 229.4; DB 4; Length 780;

Best Local Similarity 63.6%; Pred. No. 1.3e-51;

Matches 391; Conservative 0; Mismatches 206; Indels 18; Gaps 2;

QY 212 GAATTTGTTTAAAGATACAGAGGCCATTCAAGCTAAGGCTGGAAGATACAAAAACCCCT 271  
Db 10 GAGTTCTTCGTTGGGAAGCCGAGGTCCTGACGCCAAGGCTGGACACTATCAAAAGCCA 69  
QY 272 GAAGAAACCCCTTCTCCAGAAATTTACGACCATCAATTTGTCATCTTACTCCTTC 331  
Db 70 GAGATGTTCCATTTCTTCTCAAGATCTTCCACCACTCTTCTTCTTCAAAAGTGC 129  
QY 332 AAACAGTTTTTGCAATCCTGCTGCTTCAATTAACATTAACAAAGTCCATCTGAAAAATG 391  
Db 130 CCAAGGTTTTGACCCCTTTTGTCTTCAATTTGCTGCTGATGATCAATATGGAAGATG 189  
QY 392 TATTTCAAGAGTTACTTCCATTTGCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 451  
Db 190 TATTTCAATGAATGCTTACCAATTTTCACTGCGGATGGCGATGATGCGCAACTATGCA 249  
QY 452 ACTGCACTAAAGCACTTTTCAATTTATTTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 511  
Db 250 ACAGTTGCATTAGATTTTGCATGCTTGCAGGCTCTCTCAAGAGAAATTCATTTGCGCAA 309  
QY 512 TTTGTAGCTAGGTGAAATTCAGGGATGCTTCTCAAGACTACAGGCTTTTAAATTCAGCT 571  
Db 310 TATGTTGCTGAGTGAATTTCAAGAGCGGCTTCAAGATTTATAGCCCAACCAATACGCT 369  
QY 572 AAGGATAAAGAGGATTTGATGAATTTGTCATTTTCAAGGCTTCAAGGCTTCAAGAGCGGTGAG 631  
Db 370 AAGCACTAATGCTCTGATGAATTTACTTAACTTCAAGGCTTCAAGAGGATTTCAAGAGG 429  
QY 632 AAGAGAGTTGAAAGAGGCTGTGTTTGGCGAGGAAGTGAATCT---TAAAGTGTAT 688  
Db 430 AAGAGAGTAGAGAAGAGCAAGGATATTTGGACAGATGTCACCTCTGGAGGACAGTGA 489  
QY 689 GACAATGCAATG-----AAAACCGTAAATTTGATCCATCAGTGGCTTCT 733  
Db 490 GGCAGCAGATGGTGTGATGCTGTGACAGTCACTGTAAAGTTGATCCAAAGTCTTTCT 549  
QY 734 AGCTTGTACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAAGTGTGAGTACCTTTG 793  
Db 550 AAGCTATATGATGTGGGTAAATGCTTTTAAAGAGGATTTGAAAGTGAATATCTTCTC 609  
QY 794 CGCGCTCTAGACTGA 808  
Db 610 CGCGGCTTTGACTGA 624

## RESULT 10

US-09-610-040-4

; Sequence 4, Application US/09610040

; Patent No. 6465217

; GENERAL INFORMATION:

; APPLICANT: Boyes, Douglas

; APPLICANT: Davis, Keith

; APPLICANT: Woessner, Jeffrey

; APPLICANT: Gorlach, Jorn

; APPLICANT: Hamilton, Carol

; APPLICANT: Hoffman, Neil

; APPLICANT: Kloti, Andreas

; APPLICANT: Zayed, Adel

; APPLICANT: Ascenzi, Robert

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE  
; TITLE OF INVENTION: CHORISMATE MUTASE EXPRESSION OR ACTIVITY IN PLANTS





QY 798 GTCTAGACTGAAG 810  
| | | | |  
Db 135 GACTGGATTAATG 123

## RESULT 12

US-09-454-279-7  
; Sequence 7, Application US/09454279  
; Patent No. 6627798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 579  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (417)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (484)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (536)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (551)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (585)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (577)  
US-09-454-279-7

Query Match 20.2%; Score 206.2; DB 4; Length 579;  
Best Local Similarity 65.1%; Pred. No. 1.7e-45;  
Matches 304; Conservative 0; Mismatches 163; Indels 0; Gaps 0;  
QY 212 GAATTTGTTGTTAAGAAATACAGAGGCCATTCAAGCTAAGCTTGAAGATACAAAACCCCT 271  
| | | | |  
Db 3 GAGTTCCTGTTCCGGAAGCGGAGGTCTGCGACGCCAAGCTGGACACTATCAAAAGCCA 62  
| | | | |  
QY 272 GAAGAAAACGCTTCTCCAGAAATTTACCAACCATCAATTGTGCCATCTTACTCCTTC 331  
| | | | |  
Db 63 GAAGATGTTCCATCTCTCTCAAGATCTTCCACCACTGTCTTCTTCCACCAAGTGGC 122  
| | | | |  
QY 332 AAACAGTTTTTGGCATCCTGTGTGCTTCAATTAAACATAAACAAGTCCATCTGGAAAATG 391  
| | | | |  
Db 123 CCAAAGTTTTTGACCCCTTTGCTTCATTGCTGTGTAATGATGCAATATGGAAGATG 182  
| | | | |  
QY 392 TATTTCAAGAGTACTTCTCACTGCTTGTACTTCGGGTGATGGCAACTATGGCAA 451  
| | | | |  
Db 183 TATTTCAAGTAATGTCTACCATTTTCACTCGGATGGCGATGATGGCAACTATGCGAA 242  
| | | | |  
QY 452 ACTGCAAGCTAATGACCTTTTCATTATTTGCAGTCCATCTCTAGAAGGATTCACATATGGAAG 511  
| | | | |  
Db 243 ACAGTTGCATTAGATTTTCATGCTGTGAGGCTCTCTCAAGAAGAATTCATTGTGGCAA 302  
| | | | |  
QY 512 TTTGTAGCTGAGGTGAAATTCAGGATGCTCTCAAGACTACAGGCTTTAATTCAGCT 571  
| | | | |  
Db 303 TATGTTGCTGAGGTGAAATTCAAAGACGCGCTCTCAAGATTTATAGCCCAACCAATACGTGCT 362  
| | | | |

QY 572 AAGGATTAAGAGAGGATTGATGAATAATTTTGACATTTTACAAGCGTTGAAGAGCGGTGAGG 631  
| | | | |  
Db 363 AAGGACACTAATGCTCTGATGACCTTACTAACGTTTACGGCTGTTGAAGAAAAGGTCAAG 422  
| | | | |  
QY 632 AAGAGAGTTGAAAAGAGGCTGTGTTGTTGGCAGGAAGTGAATCT 678  
| | | | |  
Db 423 AAGAGAGTAGAGAAGAGGCAAGGATATTTGGACAGATGTCACCTCT 469  
| | | | |

## RESULT 13

US-09-610-040-6  
; Sequence 6, Application US/09610040  
; Patent No. 6465217  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Goriach, Jorn  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Kloti, Andreas  
; APPLICANT: Ascenzi, Robert  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE  
; FILE REFERENCE: 9128.14  
; CURRENT APPLICATION NUMBER: US/09/610,040  
; CURRENT FILING DATE: 2000-07-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 6  
; LENGTH: 5176  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-610-040-6

Query Match 9.7%; Score 98.8; DB 4; Length 5176;  
Best Local Similarity 65.1%; Pred. No. 1.3e-16;  
Matches 164; Conservative 0; Mismatches 82; Indels 6; Gaps 1;  
QY 564 TTCGAGCTAAGGATTAAGAGGATTGATGAATTTTGACATTTACAAGCGTTGAAGAGA 623  
| | | | |  
Db 3869 TTCGCTGTAGGATAGAGAGGCTTTGATGAAGCTGTGACGTTTGAGAAAGTAGAAGAA 3928  
| | | | |  
QY 624 CGGTGAGGAAGAGAGTTGAAAAGAGGCTGTGTTTGGCAGGAAGTGAATCTTTAA-- 681  
| | | | |  
Db 3929 TGGTTAAGAGAGAGAGTGCAGAGAAAGCAAGAAACGTTTGGACAAGATAAAATTCAACT 3988  
| | | | |  
QY 582 ----CAGTGTGACATGACAAATGAACCGTAAATTTGATCCATCAGTGGCTTCTAGCT 737  
| | | | |  
Db 3989 CTGCTATGCGGATGAGAGTAAGAAGATTAAGAATGAGATTCATTCGCTTCCCTTCGCA 4048  
| | | | |  
QY 738 TGTAACAAAAATTTGGTGATACCTCTCACCAGAGAGTTTCAAGTTGAGTACCTCTTCGGCC 797  
| | | | |  
Db 4049 TCTACGGGAGATGCTTATCCCTCTCACTAAGCTCGTTGAGTTGAGTATCTTCTACGTC 4108  
| | | | |  
QY 798 GTCTAGACTGAA 809  
| | | | |  
Db 4109 GTCTCGATTGAA 4120  
| | | | |

## RESULT 14

US-09-610-040-5  
; Sequence 5, Application US/09610040  
; Patent No. 6465217  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Goriach, Jorn  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffman, Neil

APPLICANT: Kloti, Andreas  
APPLICANT: Zayed, Adel  
APPLICANT: Ascenzi, Robert  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS  
FILE REFERENCE: 9128.14  
CURRENT APPLICATION NUMBER: US/09/610,040  
CURRENT FILING DATE: 2000-07-05  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 5  
LENGTH: 5099  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (4049)..(4049)  
OTHER INFORMATION: "n" indicates any nucleotide (A, C, T, or G).  
US-09-610-040-5

Query Match 7.2%; Score 73.6; DB 4; Length 5099;  
Best Local Similarity 62.5%; Pred. No. 6.2e-10;  
Matches 115; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
QY 68 ACGCTAGCTTCTGTGAGAGAGATTGGTTAGGCAAGAGATACCATCATTTATGGTCTC 127  
Db 3428 ACTCTGAGGTATAGAACTCTTTGATCCGTCAAGAGACACATTATTTGGGCTA 3487  
QY 128 ATTGAGAGAGCAAGTTCCTTAGCAATTCACACCTATGATGAAAAGTATGCTCAAATC 187  
Db 3488 TTGGAGAGAGCAAGTACTGTTACATGCTGATCTATGATCTCTGCTTTTGACATG 3547  
QY 188 CAGGGTTTCTGGCTCATTTGGTGAATTTGTTTAAAGATACAGAGGCCCATTCAGCT 247  
Db 3548 GATGGTTCAATGGTCTTTGGTTGAGTACATGTTAAAGCACTGAGAAGCTTCAGCT 3607  
QY 248 AAGG 251  
Db 3608 AAGG 3611

RESULT 15  
US-09-454-279-1  
Sequence 1, Application US/09454279  
Patent No. 6627798  
GENERAL INFORMATION:  
APPLICANT: Falco, S. Carl  
APPLICANT: Famodu, Omolayo O.  
APPLICANT: Lee, Jian-Ming  
TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
FILE REFERENCE: BB1299 US NA  
CURRENT APPLICATION NUMBER: US/09/454,279  
CURRENT FILING DATE: 1999-12-03  
EARLIER APPLICATION NUMBER: 60/110,845  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 1  
LENGTH: 552  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (442)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (456)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (460)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (472)

FEATURE:  
NAME/KEY: unsure  
LOCATION: (486)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (494)..(495)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (513)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (518)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (522)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (534)  
US-09-454-279-1  
Query Match 6.8%; Score 69.2; DB 4; Length 552;  
Best Local Similarity 55.4%; Pred. No. 4.2e-09;  
Matches 134; Conservative 0; Mismatches 108; Indels 0; Gaps 0;  
QY 2 CACGAGAGCTTAATTTCAAGATGGCCAAAGCAGCAGACAAAGTCTCTGATTCTGGGAT 61  
Db 189 CGCGCGGCCAAACAACCTCCGCAACGCCCATAGCTAAGGTAGAGAGGGTTGATCGAAGTGAC 248  
QY 62 GTGTACACGCTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTAT 121  
Db 249 ATATTGACATTGGATAGCATCAGACAAGTTTGTATTAGACTAGAAGACAGCATCATATTT 308  
QY 122 GGTCTCATTTGAGAGAGCCAAAGTTCCTAGCAATTTCTCACACCTATGATGAAAAGTATGCT 181  
Db 309 GGCCCTTTTGGAGAGAGCACAGTTTGTTCACAAATGCTGATACATATGATAGCAATGCTTTC 368  
QY 182 CAAATCCAGGGTTTGTGGCTCATTTGGTGAATTTGTTTGAATAACAGAGGCCCAT 241  
Db 369 CACATGGATGGTTTGGAGGATCTTGTGATATAGTTAGAGAACTGAGAAGCTCCATG 428  
QY 242 CA 243  
Db 429 CA 430

Search completed: September 25, 2004, 01:26:32  
Job time : 102 secs

*This page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2004, 00:33:15 ; Search time 554 Seconds  
(without alignments)  
9325.358 Million cell updates/sec

Title: US-10-624-061-15

Perfect score: 1020  
Sequence: 1 GCACGAGAGCTTAATTTCAAGAAATGCCAAAGCAGCAGCAAAAGTCTGATTCGGAA 60

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1020	100.0	1020	9	US-09-454-279-15
2	1020	100.0	1020	16	US-10-624-061-15
3	985	96.6	1079	13	US-10-425-114-7639
4	940.4	92.2	1653	13	US-10-424-599-93036
5	689.4	67.6	992	13	US-10-425-114-15059
6	465.6	45.6	525	9	US-09-454-279-5
7	465.6	45.6	525	16	US-10-624-061-5
8	327.2	32.1	1006	15	US-10-267-763-3
9	327.2	32.1	1006	15	US-10-267-763-9
10	326.2	32.0	798	11	US-09-938-842A-972
11	326.2	32.0	798	11	US-09-938-842A-972
12	260.4	25.5	694	13	US-10-424-599-141578
13	257.4	25.2	1016	17	US-10-437-963-93956
14	257.4	25.2	1231	9	US-09-454-279-13

15	257.4	25.2	1231	16	US-10-624-061-13	Sequence 13, Appl
16	254.8	25.0	1207	15	US-10-267-763-2	Sequence 2, Appl
17	254.8	25.0	1207	15	US-10-267-763-8	Sequence 8, Appl
18	254.8	25.0	1423	13	US-10-424-599-123820	Sequence 123820,
19	251	24.6	1005	9	US-09-938-842A-737	Sequence 737, App
20	250	24.5	1146	13	US-10-425-114-7285	Sequence 7285, Ap
21	239.4	23.5	1223	9	US-09-454-279-11	Sequence 11, Appl
22	239.4	23.5	1223	16	US-10-624-061-11	Sequence 27576, A
23	237.8	23.3	1197	13	US-10-425-114-27576	Sequence 14353, A
24	237.8	23.3	1200	17	US-10-767-701-14353	Sequence 49372, A
25	237.2	22.8	788	17	US-10-437-963-49372	Sequence 17, Appl
26	232.4	22.5	780	9	US-09-454-279-17	Sequence 17, Appl
27	229.4	22.5	780	16	US-10-624-061-17	Sequence 81826, A
28	227.4	22.3	427	13	US-10-424-599-81826	Sequence 4, Appl
29	227.4	22.3	1217	15	US-10-267-763-4	Sequence 10, Appl
30	211.4	20.7	1217	15	US-10-267-763-10	Sequence 7, Appl
31	211.4	20.7	579	9	US-09-454-279-7	Sequence 7, Appl
32	206.2	20.2	579	16	US-10-624-061-7	Sequence 5016, Ap
33	206.2	20.2	524	17	US-10-767-795-5016	Sequence 1332, Ap
34	204.6	20.1	1066	13	US-10-425-114-1332	Sequence 702, App
35	203.8	20.0	1469	17	US-10-437-963-702	Sequence 1370, Ap
36	201	19.7	173	17	US-10-767-701-1370	Sequence 4913, Ap
37	189.2	18.5	676	17	US-10-021-323-4913	Sequence 12291, A
38	173	17.0	622	17	US-10-021-323-12291	Sequence 3898, Ap
39	163.6	16.0	580	17	US-10-021-323-12291	Sequence 66073, A
40	160	15.7	880	13	US-10-425-114-3898	Sequence 12382, A
41	152.6	15.0	352	13	US-10-424-599-66073	Sequence 83557, A
42	143.6	14.1	542	17	US-10-021-323-12382	Sequence 101720,
43	134.6	13.2	2154	17	US-10-437-963-83557	Sequence 5596, Ap
44	132	12.9	177	13	US-10-424-599-101720	
45	129.2	12.7	600	17	US-10-767-795-5596	

## ALIGNMENTS

### RESULT 1

US-09-454-279-15  
; Sequence 15, Application US/09454279  
; Publication No. US20020184658A1  
; GENERAL INFORMATION:  
; APPLICANT: Farnco, S. Carl  
; APPLICANT: Farnco, S. Carl  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454, 279  
; EARLIER FILING DATE: 1999-12-03  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 15  
; LENGTH: 1020  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-454-279-15

Query Match	100.0%	Score 1020;	DB 9;	Length 1020;
Best Local Similarity	100.0%	Pred. No. 2.7e-264;		
Matches 1020;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GCACGAGAGCTTAATTTCAAGAAATGCCAAAGCAGCAGCAAAAGTCTGATTCGGAA	60	
Db	1	GCACGAGAGCTTAATTTCAAGAAATGCCAAAGCAGCAGCAAAAGTCTGATTCGGAA	60	
QY	61	TGTGTACACGCTAGCTTCTGTGAGAGAGATTGGTTAGGCAAGAGATACATCATTTA	120	
Db	61	TGTGTACACGCTAGCTTCTGTGAGAGAGATTGGTTAGGCAAGAGATACATCATTTA	120	
QY	121	TGGTCTTCTTACGAGAGAGCAAGTTCCTAGCAATTTCTCACACCTATGATGAAAGATGTC	180	
Db	121	TGGTCTTCTTACGAGAGAGCAAGTTCCTAGCAATTTCTCACACCTATGATGAAAGATGTC	180	

181 TCAATCCAGGTTTTTGTGGCTCATTTGGTGAATTTGTTGTAAAGATACAGAGGCCAT 240  
181 TCAATCCAGGTTTTTGTGGCTCATTTGGTGAATTTGTTGTAAAGATACAGAGGCCAT 240  
241 TCAAGCTAGGCTGGAGATACAAACCCCTGAAGAAACCGCTTCTCCAGAAATTT 300  
241 TCAAGCTAGGCTGGAGATACAAACCCCTGAAGAAACCGCTTCTCCAGAAATTT 300  
301 ACCACCATCAATTTGTGCCATCTTACTCCTTCAACAGATTTTTCATCCTGCTGCTTC 360  
301 ACCACCATCAATTTGTGCCATCTTACTCCTTCAACAGATTTTTCATCCTGCTGCTTC 360  
361 AATTACATAAACAGTCCATCTGGAATATTTTCAAGAGTTACTTCCATTTGCTTGC 420  
361 AATTACATAAACAGTCCATCTGGAATATTTTCAAGAGTTACTTCCATTTGCTTGC 420  
421 TACTTCGGGTGATGATGGCAATATATGCGCAAACTGAGCTAATGACCTTTCAATTTGCA 480  
421 TACTTCGGGTGATGATGGCAATATATGCGCAAACTGAGCTAATGACCTTTCAATTTGCA 480  
481 GTCCATCTTGAAGGATTCATATGGAAGTTTGTAGCTGAGGTGAATTCAGGATGC 540  
481 GTCCATCTTGAAGGATTCATATGGAAGTTTGTAGCTGAGGTGAATTCAGGATGC 540  
541 TCCTCAGACTAGAGCCTTTAATTCGAGTAAAGGATTAAGAGGATTCATGAAATTTGT 600  
541 TCCTCAGACTAGAGCCTTTAATTCGAGTAAAGGATTAAGAGGATTCATGAAATTTGT 600  
601 GACATTTACAAGCGTTGAAGAGCGGTGAGGAGAGATTTGAAGAGCGTGTGGTGT 660  
601 GACATTTACAAGCGTTGAAGAGCGGTGAGGAGAGATTTGAAGAGCGTGTGGTGT 660  
661 TGGCAGGAGTGAATTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
661 TGGCAGGAGTGAATTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCA 780  
721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCA 780  
781 TGAGTACCTCTTGGCGGCTAGACTGAGGAGGATTAATGATGATGATGATGATGATGAT 840  
781 TGAGTACCTCTTGGCGGCTAGACTGAGGAGGATTAATGATGATGATGATGATGATGAT 840  
841 TGAAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
841 TGAAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
901 CAATCTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
901 CAATCTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
961 TTTTCATCGCTTTTATAGGAAAGGATTAATGATGATGATGATGATGATGATGATGATGAT 1020  
961 TTTTCATCGCTTTTATAGGAAAGGATTAATGATGATGATGATGATGATGATGATGATGAT 1020

## RESULT 2

US-10-624-061-15  
; Sequence 15, Application US/10624061  
; Publication No. US20040019929A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Farnley, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/10/624,061  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: US/09/454,279  
; PRIOR FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97

; SEQ ID NO 15  
; LENGTH: 1020

; TYPE: DNA  
; ORGANISM: Glycine max

US-10-624-061-15

Query Match 100.0%; Score 1020; DB 16; Length 1020;

Best Local Similarity 100.0%; Pred. No. 2,7e-264;

Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGAGCTTAAATTTCAAGAAATGGCCAAAGCAGCAGAAACAAAGTCTGATTTCTGGGAA 60  
DB 1 GCACGAGAGCTTAAATTTCAAGAAATGGCCAAAGCAGCAGAAACAAAGTCTGATTTCTGGGAA 60  
QY 61 TGTGTACACGCTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGATACCATCTTTA 120  
DB 61 TGTGTACACGCTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGATACCATCTTTA 120  
QY 121 TGGTCTCATTTGAGAGAGCCCAAGTTCCCTAGCAATTTCTCACACCTATGATGAAAGTATGC 180  
DB 121 TGGTCTCATTTGAGAGAGCCCAAGTTCCCTAGCAATTTCTCACACCTATGATGAAAGTATGC 180  
QY 181 TCAAAATCCAGGTTTTTTGTGGCTCATTTGGTGGAAATTTGTTTGAAGATACAGAGGCCAT 240  
DB 181 TCAAAATCCAGGTTTTTTGTGGCTCATTTGGTGGAAATTTGTTTGAAGATACAGAGGCCAT 240  
QY 241 TCAAGCTAAGGCTGGAAGATACAAAACCCCTGAAGAAACCGCTTCTCCAGAAATTT 300  
DB 241 TCAAGCTAAGGCTGGAAGATACAAAACCCCTGAAGAAACCGCTTCTCCAGAAATTT 300  
QY 301 ACCACCATCAATTTGTGCCATCTTACTCCTTCAACAGATTTTTCATCCTGCTGCTTC 360  
DB 301 ACCACCATCAATTTGTGCCATCTTACTCCTTCAACAGATTTTTCATCCTGCTGCTTC 360  
QY 361 AATTAAATAAACAGTCCATCTGGAATATTTTCAAGAGTTACTTCCATTTGCTTGC 420  
DB 361 AATTAAATAAACAGTCCATCTGGAATATTTTCAAGAGTTACTTCCATTTGCTTGC 420  
QY 421 TACTTCGGGTGATGATGGCAATATATGCGCAAACTGAGCTAATGACCTTTCAATTTGCA 480  
DB 421 TACTTCGGGTGATGATGGCAATATATGCGCAAACTGAGCTAATGACCTTTCAATTTGCA 480  
QY 481 GTCCATCTTGAAGGATTCATATGGAAGTTTGTAGCTGAGGTGAATTCAGGATGC 540  
DB 481 GTCCATCTTGAAGGATTCATATGGAAGTTTGTAGCTGAGGTGAATTCAGGATGC 540  
QY 541 TCCTCAGACTAGAGCCTTTAATTCGAGTAAAGGATTAAGAGGATTCATGAAATTTGT 600  
DB 541 TCCTCAGACTAGAGCCTTTAATTCGAGTAAAGGATTAAGAGGATTCATGAAATTTGT 600  
QY 601 GACATTTACAAGCGTTGAAGAGCGGTGAGGAGAGATTTGAAGAGCGTGTGGTGT 660  
DB 601 GACATTTACAAGCGTTGAAGAGCGGTGAGGAGAGATTTGAAGAGCGTGTGGTGT 660  
QY 661 TGGCAGGAGTGAATTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
DB 661 TGGCAGGAGTGAATTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
QY 721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCA 780  
DB 721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCA 780  
QY 781 TGAGTACCTCTTGGCGGCTAGACTGAGGAGGATTAATGATGATGATGATGATGATGAT 840  
DB 781 TGAGTACCTCTTGGCGGCTAGACTGAGGAGGATTAATGATGATGATGATGATGATGAT 840  
QY 841 TGAAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 841 TGAAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900

QY 901 CAACTTCATGTAGCTGTGATTCAGAAATATTAATTTGTAGCTATAGATTGTTACCTTTAT 960  
DB 901 CAACTTCATGTAGCTGTGATTCAGAAATATTAATTTGTAGCTATAGATTGTTACCTTTAT 960  
QY 961 TTTTCATCGCTTTATTAGGAAAGCATTATATATCATGATCTTCAAAAAA 1020  
DB 961 TTTTCATCGCTTTATTAGGAAAGCATTATATATCATGATCTTCAAAAAA 1020  
RESULT 3  
US-10-425-114-7639  
; Sequence 7639, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 7639  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700677130\_FLI  
US-10-425-114-7639

Query Match 96.6%; Score 985; DB 13; Length 1079;  
Best Local Similarity 100.0%; Pred. No. 7.9e-255;  
Matches 985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 AGCTTAATTTCAAGATGCGCAAGCAGCAGCAACAAAGTCTGATTTCTGGGAATGTGTAC 67  
DB 95 AGCTTAATTTCAAGATGCGCAAGCAGCAGCAACAAAGTCTGATTTCTGGGAATGTGTAC 154  
QY 68 AGCTTAGCTTCTGTAGAGAGGATTTGGTTAGGCAAGAGATACCAATCAATTTAGTCTC 127  
DB 155 AGCTTAGCTTCTGTAGAGAGGATTTGGTTAGGCAAGAGATACCAATCAATTTAGTCTC 214  
QY 128 ATTGAGAGAGCAAGTTCCTTAGCAATTTCTCACACCTATGATGAAAGTATGCTCAATC 187  
DB 215 ATTGAGAGAGCAAGTTCCTTAGCAATTTCTCACACCTATGATGAAAGTATGCTCAATC 274  
QY 188 CAGGTTTTTTGTGCTCATTTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAGCT 247  
DB 275 CAGGTTTTTTGTGCTCATTTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAGCT 334  
QY 248 AAGCTGGAGATACAAACCCCTGAAGAAACCCCTTCTCCAGAAATTTACACCA 307  
DB 335 AAGCTGGAGATACAAACCCCTGAAGAAACCCCTTCTCCAGAAATTTACACCA 394  
QY 308 TCAATTGTGCTTCTTACTCTCTCAACAGTTTTTGGTATCTCTGGTGTCTTCAATTAAC 367  
DB 395 TCAATTGTGCTTCTTACTCTCTCAACAGTTTTTGGTATCTCTGGTGTCTTCAATTAAC 454  
QY 368 ATAAACAGTCCATCTCGAAAAATGATTTCAAGAGATTAATTCATTTGCTTGTCTTCG 427  
DB 455 ATAAACAGTCCATCTCGAAAAATGATTTCAAGAGATTAATTCATTTGCTTGTCTTCG 514  
QY 428 GGTATGATGCAACTATGCGCAACTGACCTTAAGCACTTTTCATTTATGAGTCCATC 487  
DB 515 GGTATGATGCAACTATGCGCAACTGACCTTAAGCACTTTTCATTTATGAGTCCATC 574  
QY 488 TCTAGAGAGATTCATCTATGAAAGTTTTGTAGCTGAGGTGAAATTCAGGAGTGTCTCTCAA 547  
DB 575 TCTAGAGAGATTCATCTATGAAAGTTTTGTAGCTGAGGTGAAATTCAGGAGTGTCTCTCAA 634

QY 548 GACTACGAGCTTTAATTCGAGCTAAGGATAAAGAGGATTCATGAAATTTTGACATTT 607  
DB 635 GACTACGAGCTTTAATTCGAGCTAAGGATAAAGAGGATTCATGAAATTTTGACATTT 694  
QY 608 ACAAGCGTTGAAGAGACGCTGAGGAAGAGAGTTGAAAAAGAGCTCTGTGTTTGGGAG 667  
DB 695 ACAAGCGTTGAAGAGACGCTGAGGAAGAGAGTTGAAAAAGAGCTCTGTGTTTGGGAG 754  
QY 668 GAAGTGAATCTTAACAGATGATGACATGACATGAAACCGTAAATTTGATCCATCAGTG 727  
DB 755 GAAGTGAATCTTAACAGATGATGACATGACATGAAACCGTAAATTTGATCCATCAGTG 814  
QY 728 GCTTCTAGCTTGTACAAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGTTGAGTAC 787  
DB 815 GCTTCTAGCTTGTACAAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGTTGAGTAC 874  
QY 788 CTCTTGGCGCTTAGACTGAAGGCATTACATGCAATGAGTATTAGAAATGGAAGA 947  
DB 875 CTCTTGGCGCTTAGACTGAAGGCATTACATGCAATGAGTATTAGAAATGGAAGA 934  
QY 848 TGAATATGATTTGTTGTTTCAATGATTAAGCTCTTAAGTGTATCCCTTTATTCCAACTTC 907  
DB 935 TGAATATGATTTGTTGTTTCAATGATTAAGCTCTTAAGTGTATCCCTTTATTCCAACTTC 994  
QY 908 ATGTAGCTTGTGATTCAGAAATTAATTTGTAGCTATAGATTGTTACCTTTATTTTCATC 967  
DB 995 ATGTAGCTTGTGATTCAGAAATTAATTTGTAGCTATAGATTGTTACCTTTATTTTCATC 1054  
QY 968 GCGTTTTATTAGGAAAGGCAATTT 992  
DB 1055 GCGTTTTATTAGGAAAGGCAATTT 1079

RESULT 4  
US-10-424-599-93036  
; Sequence 93036, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 93036  
; LENGTH: 1653  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_55025C.1  
US-10-424-599-93036

Query Match 92.2%; Score 940.4; DB 13; Length 1653;  
Best Local Similarity 96.6%; Pred. No. 1.1e-242;  
Matches 972; Conservative 0; Mismatches 31; Indels 3; Gaps 1;  
QY 8 AGCTTAATTTCAAGAAATGCGCAAGCAGCAGCAACAAAGTCTGATTTCTGGGAATGTGTAC 67  
DB 432 AGCTTAATTTCAAGAAATGCGCAAGCAGCAGCAACAAAGTCTGATTTCTGGGAATGTGTAC 491  
QY 68 AGCTTAGCTTCTGTAGAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGCTCTC 127  
DB 492 AGCTTAGCTTCTGTAGAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGCTCTC 551  
QY 128 ATTGAGAGAGCAAGTTCCTTAGCAATTTCTCACACCTATGATGAAAGTATGCTCAATC 187  
DB 552 ATTGAGAGAGCAAGTTCCTTAGCAATTTCTCACACCTATGATGAAAGTATGCTCAATC 611

188 CAGGGTTTTTGGGCTCAATGGTGGAAATTTGTTTAAAGATACAGAGCCCAATCAAGCT 247  
Db  
612 CAGGGTTTTTGGGCTCAATGGTGGAAATTTGTTTAAAGATACAGAGCCCAATCAAGCT 671  
Qy  
248 AAGGCTGGAGATACAAAACCTGAAGAAACCGCTCTTCCAGAAATTTACCACCA 307  
Db  
672 AAGGCTGGAGATACAAAACCTGAAGAAACCGCTCTTCCAGAAATTTACCACCA 731  
Qy  
308 TCAATTTGGCCATCTTACTCTTCAAAAGTTTTCGATCTCTGGTCTGCTTCAATTAAC 367  
Db  
732 TCAATTTGGCCATCTTACTCTTCAAAAGTTTTCGATCTCTGGTCTGCTTCAATTAAC 791  
Qy  
368 ATAAACAAGTCCATCTGGAATATTTTCAAGAGTTTACTTCCATCTGCTTACTTCG 427  
Db  
792 ATAAACAAGTCCATCTGGAATATTTTCAAGAGTTTACTTCCATCTGCTTACTTCG 851  
Qy  
428 GGTGATGAGGCAACTATCGCAAACTGAGCTTAATGACCTTTTCAATTTGAGTCCATC 487  
Db  
852 GGTGATGAGGCAACTATCGCAAACTGAGCTTAATGACCTTTTCAATTTGAGTCCATC 911  
Qy  
488 TCTAGAAGGATTCATCATGAAAGTTTGTAGTGAAGTGAATTCAGGATGCTCTCTCAA 547  
Db  
912 TCTAGAAGGATTCATCATGAAAGTTTGTAGTGAAGTGAATTCAGGATGCTCTCTCAA 971  
Qy  
548 GACTAGAGCCCTTAATTCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 607  
Db  
972 GACTAGAGCCCTTAATTCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1031  
Qy  
608 ACAAGCGTTGAAGAGAGCGTGAAGAGAGAGTTGAAAGAGAGCGTGTGTTGGGAG 667  
Db  
1032 ACAAGCGTTGAAGAGAGCGTGAAGAGAGAGTTGAAAGAGAGCGTGTGTTGGGAG 1091  
Qy  
668 GAAGTGAATCTTAACAGTGAATGACATGACATGACATGACATGACATGACATGACATG 727  
Db  
1092 GAAGTGAATCTTAACAGTGAATGACATGACATGACATGACATGACATGACATGACATG 1151  
Qy  
728 GCTTCTAGCTTTACAAAATTTGGGTGATACCTCTCACCAGAGAGTTGAGTTGAGTAC 787  
Db  
1152 GCTTCTAGCTTTACAAAATTTGGGTGATACCTCTCACCAGAGAGTTGAGTTGAGTAC 1211  
Qy  
788 CTCTTGGCGCTTAGACTGAAGGCAATTAACATGACATGACATGACATGACATGACATG 847  
Db  
1212 CTATTTGGCGCTTAGACTGAAGGCAATTAACATGACATGACATGACATGACATGACATG 1271  
Qy  
848 TGAATATGATTTGTTTCAAAATTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAA 907  
Db  
1272 TGAATATGATTTGTTTCAAAATTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAA 1331  
Qy  
908 ATGTAGCTTTGATTCAGAAATTTATTTGATGCTATGATTTGATGCTATGATTTGATG 967  
Db  
1332 ATGTAGCTTTGATTCAGAAATTTATTTGATGCTATGATTTGATGCTATGATTTGATG 1391  
Qy  
968 GCGTTTATT--AGGAAAGGCAATTTATATCATCATCTTCAAAA 1010  
Db  
1392 GCGTTTATTAGAGAAAGGCAATTTATATCATCATCTTCAAAA 1437

RESULT 5  
US-10-425-114-15059  
; Sequence 15059, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 15059  
; LENGTH: 992  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3050-017-B9\_FLI  
US-10-425-114-15059  
Query Match 67.6%; Score 689.4; DB 13; Length 992;  
Best Local Similarity 92.6%; Pred. No. 4e-175;  
Matches 747; Conservative 0; Mismatches 56; Indels 4; Gaps 2;  
Qy 208 GGTGGAATTTGTTTAAAGATACAGAGCCCAATCAAGCTAAGGCTGGAAGATACAAAA 267  
Db 1 GGTGGAATTTGTTTAAAGATACAGAGCCCAATCAAGCTAAGGCTGGAAGATACAAAA 60  
Qy 268 CCCTGAAGAAACGCTTCTTCCAGAAAAATTTACCAATCAATTTGTCCTTACTTC 327  
Db 61 CCCTGAAGAAACGCTTCTTCCAGAAAAATTTACCAATCAATTTGTCCTTACTTC 120  
Qy 328 CTTCAACAGCTTTTGCATCTGCTGCTTCAATTAACATAAACAAGTCCATCTGAA 387  
Db 121 CTTCAACAGCTTTTGCATCTGCTGCTTCAATTAACATAAACAAGTCCATCTGAA 180  
Qy 388 AATGATTTTCAAGAGTTTACTTCCATTTGCTTCTTCCGGTGTATGATGGCAATCTATGC 447  
Db 181 AATGATTTTCAAGAGTTTACTTCCATTTGCTTCTTCCGGTGTATGATGGCAATCTATGC 240  
Qy 448 GCAATCTGAGCTTAAGACCTTTTCAATTTGCTGCTTCTTCAAGAGGATTCACATG 507  
Db 241 ACAATCTGAGCTTAAGACCTTTTCAATTTGCTGCTTCTTCAAGAGGATTCACATG 300  
Qy 508 AAGTTTGTAGCTGAGGTGAATTTCAAGGATGCTCTCAAGACTACGAGCTTTAATTCG 567  
Db 301 AAGTTTGTAGCTGAGGTGAATTTCAAGGATGCTCTCAAGACTACGAGCTTTAATTCG 360  
Qy 568 AGCTAAGATTAAGAGGATTTGATGAAATTTGTTGACATTTAAGAGGTTGAAGAGCGGT 627  
Db 361 AGCTAAGATTAAGAGGATTTGATGAAATTTGTTGACATTTAAGAGGTTGAAGAGCGGT 420  
Qy 628 GAGGACAGAGTTTGAAGAAAGCGTGTGTTGGGAGAGTGAATCTTAAACAGTGA 687  
Db 421 GAGGACAGAGTTTGAAGAAAGCGTGTGTTGGGAGAGTGAATCTTAAACAGTGA 480  
Qy 688 TGCAATGACAAATGAAGAAACCGTAAATTTGATCCATGATGGCTTCTAGCTTTGACAAA 747  
Db 481 GGACGATGATGAGGAAACCAATTAAGTTTGTATCCATCAGTGGCTTCTAGCTTTGACAAA 540  
Qy 748 TTGGGTGATACCTCTCACCAGGAGTTTCAAGTTGAGTACCTCTTGGCGGCTTACAGCTG 807  
Db 541 TTGGGTGATACCTCTCACCAGGAGTTTCAAGTTGAGTACCTCTTGGCGGCTTACAGCTG 600  
Qy 808 AAGGCAATTAAGTCAAGTTAGAAATTTAAGAAATGAAGATGAATGATGTTGTTGTTTC 867  
Db 601 AAGGCAATTAAGTCAAGTTAGAAATTTAAGAAATGAAGATGAATGATGTTGTTGTTTC 660  
Qy 868 AATGATTAAGCTTTAAGTGAATCTTATTTGCAACTTCAATGCTGTTGATTCAGAA 927  
Db 661 AATGATTAAGCTTTAAGTGAATCTTATTTGCAACTTCAATGCTGTTGATTCAGAA 720  
Qy 928 ATATT-ATTTGTAGCTATAGATTTTACCTTTTATTTTTCATCGGCTTTATT-AGGAAA 983  
Db 721 ATATTAAATTTGATCTTTAGATTTGTTTACTTTTATTTTCACTGGCTTTATTAGAAGAAA 780  
Qy 984 GGCATTTATATATCATGATCTTCAAAA 1010  
Db 781 AAAAGGATTAATCATCACTTCAATA 807

RESULT 6  
US-09-454-279-5  
; Sequence 5, Application US/09454279





```
; Publication No. US20030077687A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jörn
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1006
; TYPE: DNA
; ORGANISM: Arabidopsis
US-10-267-763-3

Query Match      32.1%; Score 327.2; DB 15; Length 1006;
Best Local Similarity 64.2%; Pred. No. 2e-77;
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 AATGCCAAAGCAGCAGACAAAGTCTGATCTGGGAATGTGTACACGCTAGCTTCTGT 81
DB 69 AAGAGCTTCGAAATGCTGGTTCGGTTCCTCAATGTACTGAGCTTGAATTAAT 128
QY 82 GAGAGAGGATTGGTTAGCAGAGATACCATCTATTATGTCTCTCATTGAGAGCCAA 141
DB 129 CAGAGAAATCGTTGATTAGGCAAGACACCATCGTCTTCACTGATGAGAGAGCTAA 189
QY 142 GTTCCCTAGCAATCTCACACCTATGATGAAAGATGTCTCAATCCAGGGTCTTTTGG 201
DB 189 GTTCCCTAGCAATCTCACACCTATGATGAAAGATGTCTCAATCCAGGGTCTTTTGG 248
QY 202 CTCAATGGTGAATTTGTTAGCAAGATACAGAGCCATTCAAGCTAGCTTGAAGATA 261
DB 249 TTCTCTCACTGATTTTTCGTACAGAGACAGAAATCATCCAGCTTAAGGTAGGAGATA 308
QY 262 CAAAAACCTGAGAGAAAGCCCTTCTCCAGAAATTTTACCACCATCAATTTGTCCTATC 321
DB 309 TGAATACCCGAGAGAAATCTTCTCTTCCAGAAATTTTACCACCATCAATTTGTCCTATC 368
QY 322 TTACTCCTTCAACAGTTTTCAGTCTGCTGCTTCAATTAACATTAACAGTCCAT 381
DB 369 GCACAAATATCCATCGGCTTTCAGCCCTAAGGCTCTATCTGTTAAATTAACAAAT 428
QY 382 CTGAAAAATGATTTTCAAGAGTTACTTCCATTGCTGTCTTCCGCTGATGATGGCAA 441
DB 429 CTGGGATTTTACTTTAAAGATTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 488
QY 442 CTATCGCAAACTGCAGTAATAGCTTCAATTTATTTAGTCACTCTTCAAGGATTTCA 501
DB 489 CTATCCATCAACTGCTAGTATGATCTCGCTGTTTACAAAGCTCTTTCAGAGGATTTCA 548
QY 502 CTATGAAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCAAGCTACAGCTTCTTCT 561
DB 549 CTACGGTAAATTTGTAGCTGAGTCAATTCAGATGCTTCCAGAGATTACAGCTTCTTCT 608
QY 562 AATTCGAGCTAAGGATTAAGAGGATTGATGAATTTGATCAATTACAGCTTTGAAGA 621
DB 609 GATTCGGCTCAGGATAGAGGCTTTGATGAGCTGTTGACGTTTGAAGAGTGAAGA 668
QY 622 GACGTTGAGGAGAGATTCAAGAGGCTGTGGTGTGTTGGGCGAGGATCAATCTTAA 681
```

US-10-267-763-9/c  
; Sequence 9, Application US/10267763  
; Publication No. US20030077687A1  
; GENERAL INFORMATION:  
; APPLICANT: Goriach, Jörn  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Kloti, Andreas  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Zayed, Adel  
; APPLICANT: Ascenzi, Robert  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE  
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS  
; FILE REFERENCE: 2035DIV1  
; CURRENT APPLICATION NUMBER: US/10/267,763  
; CURRENT FILING DATE: 2002-10-09  
; PRIOR APPLICATION NUMBER: 09/610,040  
; PRIOR FILING DATE: 2000-07-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 1006  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-267-763-9

Query Match 32.1%; Score 327.2; DB 15; Length 1006;  
Best Local Similarity 64.2%; Pred. No. 2e-77;  
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 AATGCCAAAGCAGCAGACAAAGTCTGATCTGGGAATGTGTACACGCTAGCTTCTGT 81  
DB 938 AAGAGCTTCGAAATGCTGGTTCGGTTCCTCAATGTACTGAGCTTGAATTAAT 879  
QY 82 GAGAGAGGATTGGTTAGCAGAGATACCATCTATTATGTCTCTCATTGAGAGCCAA 141  
DB 878 CAGAGAAATCGTTGATTAGGCAAGACACCATCGTCTTCACTGATGAGAGAGCTAA 819  
QY 142 GTTCCCTAGCAATTTCTCACACCTATGATGAAAGATGTCTCAATCCAGGGTCTTTTGG 201  
DB 818 GTTCCCTAGCAATTTCTCACACCTATGATGAAAGATGTCTCAATCCAGGGTCTTTTGG 759  
QY 202 CTCAATGGTGAATTTGTTAGCAAGATACAGAGCCATTCAAGCTAGCTTGAAGATA 261  
DB 758 TTCTCTCACTGAGTTCGTTCAGAGAGACAGAAATCATCCAGCTTAAGGTAGGAGATA 699  
QY 262 CAAAAACCTGAGAGAAAGCCCTTCTCCAGAAATTTTACCACCATCAATTTGTCCTATC 321  
DB 698 TGAATACCCGAGAGAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 639  
QY 322 TTACTCCTTCAACAGTTTTCAGTCTGCTGCTTCAATTAACATTAACAGTCCAT 381  
DB 638 GCACAAATATCCATCGGCTTTCAGCCCTAAGGCTCTATCTGTTAAATTAACAAAT 579

382	CTCGAAATGTATTTCAAGAGTTACTTCATTCGTTCTGCTACTTCGGGTGATGATGC	441
383		
384		
385		
386		
387		
388		
389		
390		
391		
392		
393		
394		
395		
396		
397		
398		
399		
400		
401		
402		
403		
404		
405		
406		
407		
408		
409		
410		
411		
412		
413		
414		
415		
416		
417		
418		
419		
420		
421		
422		
423		
424		
425		
426		
427		
428		
429		
430		
431		
432		
433		
434		
435		
436		
437		
438		
439		
440		
441		
442	CTATGCGCAAACTGCAGCTTAATGACCTTTCATTATTGTCAGTCCATCTCTAGAAGGATCCA	501
443		
444		
445		
446		
447		
448		
449		
450		
451		
452		
453		
454		
455		
456		
457		
458		
459		
460		
461		
462		
463		
464		
465		
466		
467		
468		
469		
470		
471		
472		
473		
474		
475		
476		
477		
478		
479		
480		
481		
482		
483		
484		
485		
486		
487		
488		
489		
490		
491		
492		
493		
494		
495		
496		
497		
498		
499		
500		
501	CTATGGAAGTTTGTAGCTGAGGTGAATTCAGGGATGCTCCTCAAGACTACGAGCGCTTT	561
502		
503		
504		
505		
506		
507		
508		
509		
510		
511		
512		
513		
514		
515		
516		
517		
518		
519		
520		
521		
522		
523		
524		
525		
526		
527		
528		
529		
530		
531		
532		
533		
534		
535		
536		
537		
538		
539		
540		
541	</	

## RESULT 10

US-09-938-842A-972  
 ; Sequence 972, Application US/09938842A  
 ; Patent No. US20020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Krieps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE

FILE REFERENCE: SCRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A

: CURRENT AFFILIATION NUMBER: US/09/338,8  
 : CURRENT FILING DATE: 2001-08-24

;; PRIOR APPLICATION NUMBER: US 60/  
CURRENT FILING DATE: 2001-08-24

PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 972

```

; LENGTH: 798
TYPE: DNA

```

TYPE: DNA  
ORGANISM: Arabidopsis thaliana

; ORGANISM: *Arabidopsis thaliana*  
 URS-09-938-842A-972

US-09-938-842A-972

Query Match	32.0%: Score 32
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Query Match	32.0%;	score 32
Best Local Similarity	64.2%;	pred. No

Best Local Similarity 64.2%; Pred: No  
Matches 509; Conservative 0; Mismatches 0

RECEIVED  
JAN 10 1966  
U.S. DEPARTMENT OF JUSTICE  
FEDERAL BUREAU OF INVESTIGATION  
WASHINGTON, D.C. 20535

QY 22 AATGGCCAAGCAGCAGAACAAAGTCCT

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Db  
6 AAGAGTCTTCGAATCGGATTCGGGTTCTT

QY . 82 GAGAGGATTGGTTAGGCAAGAGGAT

Db 66 CAGAGAATCGTTGATTAGGCAAGAAGAC

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

QY 142 GTTCCCTAGCAATTCTCACACCTATGAT

[illegible]

## RESULT 11

US-09-938-842A-972  
; Sequence 972, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Krepes, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SAME, AND METHODS OF USE  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 972  
; LENGTH: 798  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana

US-09-938-842A-972

Query Match 32.0%; Score 326.2; DB 11; Length 798;  
Best Local Similarity 64.2%; Pred. No. 3.2e-77;  
Matches 509; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 ATGGCCAAAGCAGCAAGCTCTGATCTCTGGGAATGTACACGCTAGCTTCTGT 81  
DB |||||  
QY 6 AAGAGCTTCGAATCGGATTCGGGTTCTGGTTGTTCCTCAATGTACTGAGTCCTTAAT 65  
DB |||||  
QY 82 GAGAGAGGATTCGTTAGGCAAGAGATACCAATCAATTTATGTCTCTCAATGAGAGCCAA 141  
DB |||||  
QY 66 CAGAGATCGTTGATGAGCAAGAGACACCATCTCTTCAGCTTCATCGAGAGAGCTTA 125  
DB |||||  
QY 142 GTTCCCTAGCAATCTCACACCTATGATGAAAGATGCTCAAAATCCAGGGTTTTTGTGG 201  
DB |||||  
QY 126 GTTCCCACTCAATCTCTCTGCTTCGAGGAATCTCGTTGCTAGATTCTTGAAGTTTCTC 185  
DB |||||  
QY 202 CTCATTGGTGGATTTGTTGAAGATACAGAGCCATTCAAGCTTAAGCTTGGAGATA 261  
DB |||||  
QY 186 TTCTCTCACTGAGTTTTCGTCAGAGACAGAAATCATCCAGCTTAAGSTAGGAGATA 245  
DB |||||  
QY 262 CAAAAACCCCTGAAGAAAGCCCTTCTCCAGAAATTTTACCACCAATCAATTTGTGCCATC 321  
DB |||||  
QY 246 TGAATACCCGGAAGAGATCTTCTCTCTTCTGAGAAATCTCTCACTCGGTTTTTCTCTAC 305  
DB |||||  
QY 322 TTACTCTTCAACAGTTTTCGATCTCTGCTGCTTCAATTTACATATAACAAGTCCAT 381  
DB |||||  
QY 306 GCACAAATATCCATCGGCTTTCACCCCTAAGGCTCTATCTGTAACTTTAAACAAT 365  
DB |||||  
QY 382 CTGAAAAATATTTCAAAGAGTTACTTCCATTCGTTGCTACTTCCGGTGTGATGGCAA 441  
DB |||||  
QY 366 CTGGATATTTACTTTAAGAAATGCTTCTCTTCTGTTTGTCAAACCTGGCGATGGCAA 425  
DB |||||  
QY 442 CTATGCCAACTCGAGCTTAATGACCTTTCAATTTATGAGTCACTCTAGAGATCA 501  
DB |||||  
QY 426 CTATCCATCACTGCTGCTAGTATCTCGCTGTTTACAAAGCTCTTTTCGAGAGATCA 485  
DB |||||  
QY 502 CTATGAAAGTTGTAGCTGAGTGAATCAAGGATGCTCTCAAGACTAGGAGCTTT 561  
DB |||||  
QY 486 CTAGGTAAATTTGTAGCTGAGTCAATTCAGAGATGCTCCACAGATTAAGAGCTCG 545  
DB |||||  
QY 562 AATTCGAGCTAAGATAAAGAGATGATGAAATTTGATGAAATTTTACAGGTTGAGAA 621  
DB |||||  
QY 546 GATTCGGCTCAGGATAGAGAGCTTTGATGAAAGCTTTGACGTTTTCGAGAAAGTGAAGA 605  
DB |||||  
QY 622 GAGGTGAGCAAGAGATTGAAAGAGGCTGTGGTTTGGCGAGGAAGTCAATCTTAA 681  
DB |||||  
QY 606 AATGGTTAAGAGAGAGTGCAGAAAGAAAGCAAGAAAGCTTTTGGACAGAGATTAATCAA 665  
DB |||||  
QY 682 -----CAGTGATGACAAATGACAAATGAAACCCGTAATTTGATCCATCAGTGGCTCTAG 735  
DB |||||  
QY 666 CTCTGGCTATGGCGATGAGAGTAAAGAGATTAAGTGGATTCATTCCTGCTCTCG 725  
DB |||||  
QY 736 CTGTGACAAAATTTGGTGTATCTCTCACAGAGAGTTTCAGTTGAGTACTCTTGG 795  
DB |||||  
QY 726 CATCTACGGGAATGGCTTATCCCTCTCTCAAGAGCTGTTGAGTTGAGTACTCTTACG 785  
DB |||||  
QY 796 CCGTCTAGACTGA 808  
DB |||||  
QY 786 TCGTCTGATGA 798  
DB |||||

RESULT 12

US-10-424-599-141578

Sequence 141578, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J

APPLICANT: Kovalic, David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

US-10-424-599-141578

Sequence 93956, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53221)B

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 93956  
; LENGTH: 1016  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_92290C.1  
US-10-437-963-93956

Query Match 25.2%; Score 257.4; DB 17; Length 1016;  
Best Local Similarity 58.4%; Pred. No. 1.3e-58;  
Matches 469; Conservative 0; Mismatches 331; Indels 3; Gaps 1;

```
Qy 24 TGGCCAAAGCAGCAGACAAAGCTCTGATCTGGATGTGTACACGCTAGCTTCTGTGA 83
Db 155 TGACCAAGGAGGAGAAACAGAGGATAGATCAAGTGAATFACCTTGGACAACTTA 214
Qy 84 GAGAGGATTTGGTTAGGCAAGAGGATACCATCTATTTATGGTCTCATTTGAGAGAGCCAAAGT 143
Db 215 GAACCTCTCTGGTTAGGCAAGAGGATACCATCTATTTATGGTCTCATTTGAGAGAGCCAAAGT 274
Qy 144 TCCCTAGCAATTCACACCTATGATGAAAAGTATGCTCAATCCAGGCTTTTGTGGCT 203
Db 275 TTTGTCTCAATGCTGATATATATGATAAAATGCTTTCCATGTGGATGGATTTGTGGCT 334
Qy 204 CATTTGGTGGATTTGTTTAAAGAAATACAGAGGCCATTCAGCTAAGGCTGGAGATACA 263
Db 335 CTITGGTTGATTCATGTTAGAGAAACCGAAACCTACATCAACAGGTTGGAGATACA 394
Qy 264 AAAACCTCGAAGAAAACGCTTCTTCCGAGAAATTTACACCATCAATTTGTGCCATCTT 323
Db 395 AGAGCCCTGTAGACACCCATCTTCCGAGGATCTGCTGAACCACTTTGCCACCTC 454
Qy 324 ACTCTTTCAACACGTTTTTGCATCCTGCTGCTTCAATTAACATAAACAAGTCCATCT 383
Db 455 TCCAGTATCAAGGTTTTGCACTTATGCTGATCTATTAATATCAACAGAGATTT 514
Qy 384 GGAATAATGATTTCAAGAGTACTTCCATTTGCTGCTACTTCGGTGTATGTGGCAACT 443
Db 515 GGAATAATGATTTTGTAGAGCTTCTTCCAAAGATTTAGTGAAGAGGAAGTGTGTTAAT 574
Qy 444 ATGGCAAACTGCAGCTAATGACCTTTTCAATTTGCAAGTCCATCTCTAGAAGGATTCAC 503
Db 575 ATGGATCCAGTCTCTTTGTGACACGATCTGCTTGCAGGCGCTTCCAAAAGAAATTCAC 634
Qy 504 ATGAAAAGTTTGTAGCTGAGGTGAAATTCAGGATGCTCTCAAGACTACAGCCCTTTAA 563
Db 635 ATGTAAGTTTGTGCGCAGAGCTAAGTTTCAAGAGTCTCTGAGCTTTACATGCTGCGA 694
Qy 564 TCGAGCTAAGGATAAAGAGGATGATGAAATTTGTTGACATTTAAGGCTTTAGAGAGA 623
Db 695 TAATAGCACAGGACTGCGATCAACTAATGCACTCTCTCACTATGAACCGTGGAGCGTG 754
Qy 624 CGGTGAGGAGAGAGTTGAAAAGAGGCTGTGGTGTGTTGGCGAGGAGTCAATCTTAAACA 683
Db 755 CTATGACATAGGTTGAGCTAAGCTAAGATCTTTTGACAGAGGTTGGATTTAGGCG 814
Qy 684 GTGATGACATGACAAAT---GAAAACCGTAAATTTGATCCATCAGTGGCTTTAGTGTGT 740
Db 815 CTGAAGACACCGGCTCCACCAATGTACAAGATAAGGCCAGTTTGGTGGCTGAACGTGT 874
Qy 741 ACAAATAATGGGTGATACCTCTCAAGAGGAGGTTGAGTTGAGTACCTCTTCGGCGGTC 800
Db 875 ACAGCTACAGGATCATGCGGTAAACAGGAGGTTCAAGTAGCCCTACTTCTGAGGAGAT 934
Qy 801 TAGACTGAAGGCATTACAATGCA 823
Db 935 TGGATTGATTTTACGATTGTA 957
```

## RESULT 14

US-09-454-279-13  
; Sequence 13, Application US/09454279  
; Publication No. US20020184658A1

; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: B01299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 13  
; LENGTH: 1231  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; US-09-454-279-13

Query Match 25.2%; Score 257.4; DB 9; Length 1231;  
Best Local Similarity 58.4%; Pred. No. 1.5e-58;  
Matches 469; Conservative 0; Mismatches 331; Indels 3; Gaps 1;

```
Qy 24 TGGCCAAAGCAGCAGACAAAGTCTGATCTGGGAATGTGTACACGCTAGCTTCTGTGA 83
Db 239 TGACCAAGGAGGAGAAACAGAGGATAGATCAAAAGTGAATACTAGACCTTGGACAACTTA 298
Qy 84 GAGAGGATTTGGTTAGGCAAGAGGATACCATCTATTTATGGTCTCATTTGAGAGAGCCAAAGT 143
Db 299 GAACCTCTCTGGTTAGGCAAGAGGATACCATCTATTTATGGTCTCTTAGAGAGAGCAGT 358
Qy 144 TCCCTAGCAATTTCTCACACCTATGATGAAAAGTATGCTCAATCCAGGCTTTTGTGGCT 203
Db 359 TTTGTCTCAATGCTGATATATATGATAAAATGCTTTCCATGTGGATGGATTTGATGGCT 418
Qy 204 CATTTGGTGGATTTGCTTAAAGATACAGAGGCCATTCAGCTAAGCTGGAAGATACA 263
Db 419 CTITGGTTGATTCATGTTAGAGAAACCGAAACCTACATCAACAGCTTGGAGATACA 478
Qy 264 AAAACCTCGAAGAAAACGCTTCTTCCAGAAAATTTTACCACCAATCAATTTGCGCATCT 323
Db 479 AGAGCCCTGTAGCAGCACCCATCTTTTCCGAGGATCTGCTGAACCACTGTTGCCACCTC 538
Qy 324 ACTCTTTCAACAGTTTTTGTGATCTCTGGTGTCTTCAATTTAAATAAACAAGTCCATCT 383
Db 539 TCCAGTATCCAAAGGTTTTGCTATCTCTTGTGATTTATTAATCAACAAGGAGATTT 598
Qy 384 GGAATAATGATTTCAAGAGTACTTCCATTTGCTGCTACTTCCGGTGTGATGATGCAACT 443
Db 599 GGAATAATGATTTTGTAGAGCTTCTTCCAGATTTAGTGAAGAGGAAAGTGTATGTAAT 658
Qy 444 ATGGCAAACTGCAGCTAATGACCTTTTCAATTTGCACTCCATCTCTAGAAGGATTCAT 503
Db 659 ATGGATCCAGTCTCTTTGTGACACGATCTGCTTGCAGGCGCTCTCCAAAAGAAATTCAC 718
Qy 504 ATGGAAGTTTGTAGCTGAGGTGAAATTCAGGATGCTCTCAAGACTACAGGCTTTAA 563
Db 719 ATGGAAGTTTGTGCGCAGGCTAAGTTTCAAGAGTCTCTTGAAGCTTACATGCTCGGA 778
Qy 564 TTCGAGCTAAGGATAAAGAGGATTTGATAAATTTTGACATTTTACAAGCGTTGAGAGA 623
Db 779 TAATAGCAGGACTGCGATCAACTAATGCACTCTCCATCTATGAACCGTGGAGCGTG 838
Qy 624 CGGTGAGGAGAGGATTTGAAAAGAGGCTGTGGTGTGTTGGCGAGGAGTGAATCTTAACA 683
Db 839 CTATTTGAACATAGGTTGGAAGCTAAGGCTTAAGCTTTTGGACAGAGGTTGATTTAGCG 898
Qy 684 GTGATGACAAATGACAAAT---GAAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTGT 740
Db 899 CTGAAGACACCGCGCTCCACCAATGTACAGATAAGGCCAGTTTGGTGGCTGAACGT 958
Qy 741 ACAAATAATGGGTGATACCTCTCAAGAGGAGGTTGAGTTGAGTACCTCTTTCGCCCTC 800
Db 959 ACAGCTACAGGATCATGCGCTTAACCAAGAGGTTTCAAGTAGCCCTACTTGTGAGGAGAT 1018
```

QY 801 TAGACTGAAGGCATTACAATGCA 823  
 Db 1019 TGGATTGATTGTTTACGATTGTA 1041

RESULT 15

US-10-624-061-13  
 ; Sequence 13, Application US/10624061  
 ; Publication No. US20040019929A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Falco, S. Carl  
 ; APPLICANT: Ramodu, Omolayo O.  
 ; APPLICANT: Lee, Jian-Ming  
 ; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 ; FILE REFERENCE: B01299 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/624,061  
 ; CURRENT FILING DATE: 2003-07-21  
 ; PRIOR APPLICATION NUMBER: US/09/454,279  
 ; PRIOR FILING DATE: 1999-12-03  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 13  
 ; LENGTH: 1231  
 ; TYPE: DNA  
 ; ORGANISM: Oryza sativa  
 ; US-10-624-061-13

Query Match 25.2%; Score 257.4; DB 16; Length 1231;  
 Best Local Similarity 58.4%; Pred. No. 1.5e-58;  
 Matches 469; Conservative 0; Mismatches 331; Indels 3; Gaps 1;

QY 24 TGGCCAAAGCAGCAGAAAGTCCTGATCTGGGAATGTGTACACGCTAGCTTCTGTGA 83  
 Db 239 TGACCAAGGAGGAGAAACACAGAGATAGATCAAAAGTGAATACTGACCTTGGACAAATTA 298  
 QY 84 GAGAGGATTGGTTAGGCAAGAGATACCATCATTTATGGTCTCATTTGAGAGAGCAAGT 143  
 Db 299 GAACCTCCTTGGTTAGGCAAGAGAGACAGCATATATTCAGCCTCTTAGAGAGAGCACAGT 358  
 QY 144 TCCTTAGCAATTCCTACACTATGATGAAAGATGTCTCAAAATCCAGGGTTTTTGTGGCT 203  
 Db 359 TTTGCTACAATGCTGATATATATGATAAAATGCTTTCCATGTGGATGGATTGTGATGCT 418  
 QY 204 CATTGGTGGAAATTTGTTTGAATACAGAGAGGCCATTCACGCTAAGGCTGGAAGATACA 263  
 Db 419 CTTTGGTTGAATTCATGTTTAGAAGAACCCGAAACTACATCAACAGGTTGGAGATACA 478  
 QY 264 AAAACCCCTGAAGAAAACGCTTTCTCCAGAAAATTTACCAACATCAATTTGCCATCTT 323  
 Db 479 AGAGCCCTGATGAGCACCATTCTTTCCGGAGGATCTGCTCAACACTTTGCCACCTC 538  
 QY 324 ACTCCTTCAACAGTTTTTTCATCTCGTGTGCTTCAATTAACATAACAGTCCATCT 383  
 Db 539 TCAGATATCCAAAGGTTTTGCAATCTGCTGATTTATTAATATCAACAGGAGATTT 598  
 QY 384 GGAATAATGTATTTCAAGAGTTACTTCCATTTGCTTACTTTCGGGTGATGATGCCAAT 443  
 Db 599 GGAATAATGTATTTGATGAGCTTCTTCAAGATTAGTGAAGAGGAAGTGAATTAAT 658  
 QY 444 ATGCGCAAACTGAGCTAATGACCTTTCAATTTGAGAGTCCATCTCTAGAGGATTTCACT 503  
 Db 659 ATGGATCCAGTGTCTTTGTGACAGCATCTGCTTGGCGGCTCTCCAAAAGAAATTCAT 718  
 QY 504 ATGGAAGTTTGTAGTGTAGGTGAATTCAGGATGCTCTCTCAAGACTACGAGCCTTTAA 563  
 Db 719 ATGGTAAATTTTGGCAGAGGCTTAGTTTCAAGAGTCTCTGAGCTTACATGCTCGGA 778  
 QY 564 TTCGAGCTAAGGATAAGAGGATTGATGAATTTGTTGACATTTTCAAGCGTTGAAGAGA 623  
 Db 779 TAATAGCAGGAGCTCGGATCAACTAATGCACTCTCTCACTATGAAACGGTGGAGCGTG 838

QY 624 CGGTGAGAAAGAGAGTTGAAAAGAGGCTGTGTGTTTGGGCAAGAGTGAATCTTTAACA 683  
 Db 839 CTATTGAACATAGGGTGAAGCTAAGGCTAAGATCTTTGGACAGAGAGGTGGATTAGGCG 898  
 QY 584 GTGATGACATGACAAAT---GAAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTTGT 740  
 Db 899 CTGAAGACAACGGCGCTCCACCAATGTACAAGATAAGGCCACGTTTGGTGGCTGAATCTGT 958  
 QY 741 ACAAAAATTGGGTGATACCTTCTCACCAGAGAGGTTTCAGGTTGAGTACCTTCTTCGCCGCTC 800  
 Db 959 ACAGCTACAGGATCATGCGCGCTAACCAAGAGGCTTCAAGTAGCCTACTTGTCTGAGGAGAT 1018  
 QY 801 TAGACTGAAGGCATTACAATGCA 823  
 Db 1019 TGGATTGATTGTTTACGATTGTA 1041

Search completed: September 25, 2004, 02:51:17  
 Job time : 557 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 23:12:46 ; Search time 3094 Seconds

(without alignments)  
9844.675 Million cell updates/sec

Title: US-10-624-061-15  
Perfect score: 1020  
Sequence: 1 gcacgagagcttaatttcaa.....atcttcaaaaaaaaaa 1020

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est6:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	541	53.0	560	10 BE821647	BE821647 GW700015A
2	525	51.5	563	10 AW666427	AW666427 sk36d12.y
3	511	50.1	512	12 EG363365	EG363365 sac19f08.
4	477	46.8	477	9 AI973566	AI973566 sc88e02.y

5	469	45.0	620	10 BF009709	BF009709 ss83b10.y
6	468.4	45.9	639	10 BE660542	BE660542 2-B12 Gna
7	457.6	44.9	648	10 AW774625	AW774625 EST333776
8	438.2	43.0	696	12 B1273174	B1273174 NF091B12F
9	431.2	42.3	654	10 AW775842	AW775842 EST334907
10	416.8	40.9	605	10 BF649665	BF649665 NF081H11E
11	415.6	40.7	422	12 BG239446	BG239446 sab70d03.
12	404	39.6	560	12 BM093460	BM093460 saj09c02.
13	400.4	39.3	529	14 CB827129	CB827129 LJNEST70F
14	388.8	38.1	605	12 B1420527	B1420527 LJNEST585
15	388.8	38.1	726	14 CA800612	CA800612 saul9a05.
16	321	31.5	524	14 CB828627	CB828627 LJNEST889
17	313.2	30.7	531	14 CB828445	CB828445 LJNEST86F
18	313.2	30.7	557	14 CB829043	CB829043 LJNEST93e
19	310.2	30.4	527	10 BE800482	BE800482 sg93h09.y
20	308.2	30.2	516	14 CB827444	CB827444 LJNEST74h
21	303.2	29.7	501	14 CB828062	CB828062 LJNEST82C
22	302.8	29.7	549	14 CB828614	CB828614 LJNEST88F
23	299.6	29.4	529	14 CB826768	CB826768 LJNEST66F
24	298	29.2	541	14 CB829238	CB829238 LJNEST95G
25	283.6	27.8	748	14 CD824727	CD824727 BN25.058J
26	271.6	26.6	466	10 AW720060	AW720060 LJNEST13b
27	262	25.7	459	12 B1419161	B1419161 LJNEST29b
28	258	25.3	885	13 BU693059	BU693059 SFD0085.P
29	251	24.6	698	10 BE577489	BE577489 L48-2335T
30	249.2	24.4	710	14 CF475712	CF475712 RTW2.11
31	249	24.4	394	10 AW774873	AW774873 EST334024
32	242.2	23.7	810	14 CD427426	CD427426 SA1.30.H0
33	241.4	23.7	427	12 B1420751	B1420751 LJNEST61c
34	239.4	23.5	1222	11 AY103806	AY103806 Zea mays
35	237.8	23.3	1196	11 A1107907	A1107907 Zea mays
36	236	23.1	660	13 BX253798	BX253798 BX253798
c 37	230.4	22.6	867	14 CB619882	CB619882 OSIIEA04L
38	229.4	22.5	680	13 CA080410	CA080410 SCVPAM105
39	227.8	22.3	821	14 CK196770	CK196770 FGAS00523
40	227.4	22.3	427	12 BG359738	BG359738 sac28a03.
c 41	216.6	21.2	725	14 CF486896	CF486896 POU1.40.B
42	213.4	20.9	696	14 CD919148	CD919148 G608.112E
c 43	212.4	20.8	607	9 AI486636	AI486636 EST244957
44	211	20.7	798	14 CB619881	CB619881 OSIIEA04L
45	210.8	20.7	623	14 CD998155	CD998155 QB88e10.x

#### ALIGNMENTS

RESULT 1  
BE821647/c  
LOCUS BE821647 560 bp mRNA linear EST 24-MAY-2001  
DEFINITION GW700015A10A11 Gm-r1070 Glycine max cDNA clone Gm-r1070-5781 3', mRNA sequence.  
ACCESSION BE821647  
VERSION BE821647.1 GI:10253981  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
REFERENCE 1 (bases 1 to 560)  
AUTHORS Vodka, L., Kelm, P., Shoemaker, R., Retzel, E., Khanna, A., Corvett, V., Erpelting, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.  
TITLE A Functional Genomics Program for Soybean (NSF 9872565)  
JOURNAL Unpublished (1999)  
COMMENT Other ESTs: AI973566 corresponding to Gm-cl019-435 (5')  
Contact: Vodka, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147





subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript II Xr Predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Kelm and Dr. Virginia Coryell."

## ORIGIN

Query Match 51.5%; Score 525; DB 10; Length 563;  
 Best Local Similarity 96.2%; Pred. No. 7.1e-108; Indels 0; Gaps 0;  
 Matches 537; Conservative 0; Mismatches 21;

Qy 94 GGTAGGCAAGAGATACCATCATTTATGGTCTCATTTGAGAGCCAAAGTTCCTAGCAA 153  
 Db 5 GGTAGGCAAGAGATACCATCATTTATGGTCTCATTTGAGAGCCAAAGTTCCTAGCAA 64

Qy 154 TTCTCACACCTATGATGAAGAATGATGCTCAATCCAGGGTTTTTGTGGCTCATTGGTGA 213  
 Db 65 TCCTCACACCTATGATGAAGAATGATGCTCAATCCAGGGTTTTTGTGGCTCATTGGTGA 124

Qy 214 ATTTGTTGTTAAGAAATACAGAGGCCATTCAGCTAAGCTGGAAGATACAAAAACCTGA 273  
 Db 125 ATTTGTTGTTAAGAAATACAGAGGCCATTCAGCTAAGCTGGAAGATACAAAAACCTGA 184

Qy 274 AGAAACGCCCTTCTCCAGAAAAATTTACACCATCAATTTGTGCCATCTTACTCTTCAA 333  
 Db 185 AGAAACGCCCTTCTCCAGAAAAATTTACACCATCAATTTGTGCCATCTTACTCTTCAA 244

Qy 334 ACAGTTTTTGATCTGTGTGCTTCAATTAATAAACAAAGTCCATCTGGAATAATGTA 393  
 Db 245 ACAGTTTTTGATCTGTGTGCTTCAATTAATTAATTAATTAATTAATTAATTAATGTA 304

Qy 394 TTTCAGAGATTAATTCATTTGCTTGTGCTTGTGCTGATGATGCGAACTATGCGAAAC 453  
 Db 305 TTTCAGAGATTAATTCATTTGCTTGTGCTTGTGCTGATGATGCGAACTATGCGAAAC 364

Qy 454 TGCAGCTAATGACCTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513  
 Db 365 TGCAGCTAATGACCTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424

Qy 514 TGTAGCTGAGTGAAATTCAGGGATGCTCTCAAGACTACGAGCTTTAATTCAGCTAA 573  
 Db 425 TGTAGCTGAGTGAAATTCAGGGATGCTCTCAAGACTACGAGCTTTAATTCAGCTAA 484

Qy 574 GGATAAGAGATGATGAATTTGTTGACATTTTACAGCGTTGAAGAGCGTGAAGAA 633  
 Db 485 GGATAAGAGATGATGAATTTGTTGACATTTTACAGCGTTGAAGAGCGTGAAGAA 544

Qy 634 GAGAGTTGAAAGAGAGGC 651  
 Db 545 AGAGTTGAAAGAGAGGC 562

## RESULT 3

BG363365 512 bp mRNA linear EST 28-NOV-2001  
 LOCUS sac19f08.v1 Gm-cl051 Glycine max cDNA clone GENOME SYSTEMS CLONE  
 DEFINITION ID: Gm-cl051-2943 5' similar to TR:Q9S7H4 Q9S7H4 CHORISMATE MUTASE  
 ; mRNA sequence.

## ACCESSION

BG363365

## VERSION

BG363365.1 GI:13252454

## KEYWORDS

Glycine max (soybean)

## SOURCE

Glycine max

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

## REFERENCE

## AUTHORS

1. (bases 1 to 512)  
 Shoemaker,R., Kelm,P., Vodkin,L., Erpelnding,J., Coryell,V.,  
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
 McCann,R., Waterston,R. and Wilson,R.

## TITLE

## JOURNAL

## COMMENT

Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

High quality sequence stop: 397.

## FEATURES

## source

1..512

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-2943"

/tissue\_type="floral meristematic mRNA"

/lab\_host="DH10B"

/clone\_lib="Gm-cl051"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The cDNA library was constructed from floral

meristematic mRNA provided by Dr. Halina Knap of Clemson

University. Complementary DNA was synthesized from mRNA

using a primer consisting of a poly(dT) sequence with a

XhoI restriction site. EcoRI adapters were ligated to the

blunt-ended cDNA fragments followed by XhoI digestion. The

cDNA fragments were directionally cloned into the

EcoRI-XhoI restriction site of the pBluescript vector. The

ligated cDNA fragments were transformed into DH10B host

cells (GibcoBRL). This library was constructed in the

laboratory of Dr. Randy Shoemaker."

## ORIGIN

Query Match 50.1%; Score 511; DB 12; Length 512;  
 Best Local Similarity 99.8%; Pred. No. 1e-104;  
 Matches 511; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 39 AACAAAGTCTGATTCCTGGGAATGTACACGCTAGCTTCTGTGAGAGAGGATTTGGTTA 98

Db 1 AACAAAGTCTGATTCCTGGGAATGTACACGCTAGCTTCTGTGAGAGAGGATTTGGTTA 60

Qy 99 GGCAAGAGGATACCATTCATTTATGGTCTCATTTGAGAGAGCCAAAGTTCCTAGCAATTC 158

Db 61 GGCAAGAGGATACCATTCATTTATGGTCTCATTTGAGAGAGCCAAAGTTCCTAGCAATTC 120

Qy 159 ACACCTTATGATGAAGAAAGTATGCTCAATCCAGGGTTTTTGTGGCTCATTGGTGGAAATTTG 218

Db 121 ACACCTTATGATGAAGAAAGTATGCTCAATCCAGGGTTTTTGTGGCTCATTGGTGGAAATTTG 180

Qy 219 TTGTTAAGAAATACAGAGGCCATTCAGCTTAAGCTGGAAGATACAAAAACCTGGAAGAAA 278

Db 181 TTGTTAAGAAATACAGAGGCCATTCAGCTTAAGCTGGAAGATACAAAAACCTGGAAGAAA 240

Qy 279 ACACCTTCTTCCAGAAATTTACCAACCATCAATTTGTCCTTACTCTCTCAACAGT 338

Db 241 ACACCTTCTTCCAGAAATTTACCAACCATCAATTTGTCCTTACTCTCTCAACAGT 300

Qy 339 TTTTTCATCTCTGCTGCTTCAATTAACATAAACAAGTCCATCTCGAAAAATGATTTC 398

Db 301 TTTTTCATCTCTGCTGCTTCAATTAACATAAACAAGTCCATCTCGAAAAATGATTTC 360

399 AAGAGTTACTTCCATTCCTGTTACTTCTCGGTGATGCAACTATGCGCAAACTGCAG 458  
 Db |||||||  
 361 AAGAGTTACTTCCATTCCTGTTACTTCTCGGTGATGCAACTATGCGCAAACTGCAG 420  
 Qy |||||||  
 459 CTAATGACCTTTCATTTACGAGTCCATCTCTAGAGGATTCATGGAAGCTTTGTAG 518  
 Db |||||||  
 421 CTAATGACCTTTCATTTACGAGTCCATCTCTAGAGGATTCATGGAAGCTTTGTAG 480  
 Qy |||||||  
 519 CTGAGTGAAATTCAGGGATGCTCCTCAAGAC 550  
 Db |||||||  
 481 CTGAGTGAAATTCAGGGATGCTCCTCAAGAC 512

RESULT 4  
 AI973566  
 LOCUS  
 DEFINITION sc88802.y1 Gm-cl019 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-cl019-435 5', similar to SW:CHMU\_ARATH\_P42738 CHORISMATE MUTASE  
 PRECURSOR ;, mRNA sequence.

ACCESSION  
 VERSION AI973566.1 GI:5770392  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max

REFERENCE  
 AUTHORS  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

1 (bases 1 to 477)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,  
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McCann, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com

Insert Length: 653 Std Error: 0.00

Seq primer: -40RP from Gibco

High quality sequence stop: 447.

Location/Qualifiers

FEATURES  
 source

1. 477

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl019-435"

/tissue\_type="Immature seed coats of greenhouse grown

plants"

/lab\_host="DH10B (Gibco BRL)"

/notes="Vector: pSPORt1 (Life Technologies); Site 1: Not I;

Site 2: Sal I; This cDNA library was constructed from mRNA

isolated from immature seed coats (200-300 mgs) of

greenhouse grown plants. The library was prepared using

the Life Technologies pSuperScript cDNA library

construction kit. Complementary DNA was synthesized from

mRNA using a poly (dT) sequence with a Not I restriction

site. Sal I linker adapters were ligated to the

blunt-ended cDNA fragments followed by Not I digestion.

The cDNA fragments were directionally cloned into the Not

I-Sal I restriction site of the pSPORt1 vector. The

ligated cDNA fragments were transformed into E.coli

ElectroMax DH10B host cells (Gibco BRL). This library was  
 constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

ORIGIN  
 Query Match 46.8%; Score 477; DB 9; Length 477;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-97;  
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 383 TGGAAATGTTATTTCAAAGAGTTACTTCCATTTGCTTCTCGGTGATGCAAC 442  
 Db 1 TGGAAATGTTATTTCAAAGAGTTACTTCCATTTGCTTCTCGGTGATGCAAC 60  
 Qy 443 TATGCGAAACTGCAGCTAATGACCTTTTATTGTCAGTCCATCTCTAGAAGATTTCAC 502  
 Db 61 TATGCGAAACTGCAGCTAATGACCTTTTATTGTCAGTCCATCTCTAGAAGATTTCAC 120  
 Qy 503 TATGGAAGTTTGTAGCTGAGGTGAAATTCAGGATCTCTCAAGACTACGAGCTTTA 562  
 Db 121 TATGGAAGTTTGTAGCTGAGGTGAAATTCAGGATCTCTCAAGACTACGAGCTTTA 180  
 Qy 563 ATTGAGCTTAAGATAAAGAGGATTGATGAATTTGTGACATTTACAGCGTTGAGAG 622  
 Db 181 ATTGAGCTTAAGATAAAGAGGATTGATGAATTTGTGACATTTACAGCGTTGAGAG 240  
 Qy 623 ACGGTGAGGAAGAGAGTTGAAAAGAGGCTCTGTGTTTGGGAGGAGGAATGAATCTTAAC 682  
 Db 241 ACGGTGAGGAAGAGAGTTGAAAAGAGGCTCTGTGTTTGGGAGGAGGAATGAATCTTAAC 300  
 Qy 683 AGTGATGACAAATGACAAATGAAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTGTAC 742  
 Db 301 AGTGATGACAAATGACAAATGAAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTGTAC 360  
 Qy 743 AAAAAATGGGTGATACCTCTCACCAGGAGGTTTCAGGTTCAGTACCTCTTGGCGCGCTCTA 802  
 Db 361 AAAAAATGGGTGATACCTCTCACCAGGAGGTTTCAGGTTCAGTACCTCTTGGCGCGCTCTA 420  
 Qy 803 GACTGAAGGCATTACAATGCAATGCAATTTAGAAATTTAGAAATTTAGAAATTTAGATGT 859  
 Db 421 GACTGAAGGCATTACAATGCAATGCAATTTAGAAATTTAGAAATTTAGAAATTTAGATGT 477

RESULT 5  
 BF009709  
 LOCUS  
 DEFINITION ss83b10.y1 Gm-cl064 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-cl064-692 5', similar to TR:Q9S7H4 Q9S7H4 CHORISMATE MUTASE ;,  
 mRNA sequence.

ACCESSION BF009709

VERSION BF009709.1 GI:10709985

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

1 (bases 1 to 620)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,  
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McCann, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
High quality sequence stop: 404.

FEATURES

source  
1. .620  
Location/Qualifiers  
/organism="Glycine max"  
/mol\_type="mRNA"  
/cultivar="Williams"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl064-692"  
/tissue\_type="seedling epicotyls"  
/dev\_stage="2 week old"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl064"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI. The cDNA library was constructed from mRNA isolated  
from the epicotyls of 2 week old seedling in a growth  
chamber, excised above the soil level, and the plants  
were placed in a 100 ppm solution of auxin for 24 hours  
prior to harvesting. Complementary DNA was synthesized  
from mRNA using a primer consisting of a poly(dT)  
sequence with a XhoI restriction site. EcoRI adapters  
were ligated to the blunt-ended cDNA fragments followed  
by XhoI digestion. The cDNA fragments were directionally  
cloned into the EcoRI-XhoI restriction site of the  
pBluescript vector. The ligated cDNA fragments were  
transformed into DH10B host cells (GibcoBRL). This  
library was constructed in the laboratory of Dr. Randy  
Shoemaker."

ORIGIN

Query Match 46.0%; Score 469; DB 10; Length 620;  
Best Local Similarity 91.7%; Pred. No. 2.9e-95;  
Matches 518; Conservative 0; Mismatches 45; Indels 2; Gaps 2;  
Qy 8 AGCTTAATTTCAAGATGGCCAAAGCAGAGACAAAGTCTGATCTGGGATGTGAC 67  
Db 57 ACCTTAATTTCAAGATGGCCAAAGCAGAGACAAAGTCTGATCTGGGATGTGAC 116  
Qy 68 ACGTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGATACCAATCATTTATGCTCTC 127  
Db 117 ACGTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGATACCAATCATTTATGCTCTC 176  
Qy 128 ATTGAGAGACCAAGTTCCTAGCAATTCACACCTATGATGAAAGATGCTCAATC 187  
Db 177 ATTGAGAGACCAAGTTCCTAGCAATTCACACCTATGATGAAAGATGCTCAATC 236  
Qy 188 CAGGCTTTTGTGCTCATTTGGTGGAAATTTGTTAAGAAATACAGAGCCATTCAGCT 247  
Db 237 CAGGCTTTTGTGCTCATTTGGTGGAAATTTGTTAAGAAATACAGAGCCATTCAGCT 296  
Qy 248 AAGCTCGAAGATACAAAAACCCCTGAAGAAAAAGCCCTTCTCCAGAAAAATTTACACCA 307  
Db 297 AAGCTCGAAGATACAAAAACCCCTGAAGAAAAAGCCCTTCTCCAGAAAAATTTACACCA 356  
Qy 308 TCAATTTGGCCATCTTACTCTTCAACACAGTTTGTGATCTGCTGCTCAATTAAC 367  
Db 357 TCAATTTGGCCATCTTACTCTTCAACACAGTTTGTGATCTGCTGCTCAATTAAC 416  
Qy 368 ATAAACAGTCCATCTGGAATATGTTTCAAGAGATTAATTCATTTGCTTACTTCG 427  
Db 417 ATTAACAGTCCATCTGGAATATGTTTCAAGAGATTAATTCATTTGCTTACTTCG 475  
Qy 428 GGTATGATGCAACTATGCAACTGCACTAATGACCTTTTCAATTTGCTGCTCATC 487  
Db 476 GGGATGATGCAACTATGCAACTGCACTAATGACCTTTTCAATTTGCTGCTCATC 535  
Qy 488 TCTA-GAAGGATTCACATGGAAGTTTGTAGCTGAGGTGAAATTCAGGATGCTCTCA 546  
Db 536 TTTAGGAGATTTCCCTTGGAAAGTTTGTAGCTGAGGTGAAATTCAGGATGCTCTCC 595  
Qy 547 AGACTAGAGCCTTTAATTCAGCT 571

Db 596 AAACCTAGAGCCTTATTCGAGCT 620

RESULT 6

BE660542  
LOCUS 2-B12 GmaxSC Glycine max cDNA, mRNA sequence.  
DEFINITION BE660542  
ACCESSION BE660542  
VERSION BE660542.1 GI:9986434  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 639)  
Harris, N., Chapman, B.P. and Gijzen, M.  
Gene expression in developing soybean seed coats  
Unpublished (2000)  
Contact: Gijzen M  
Agriculture and Agri-Food Canada  
1391 Sandford Street, London, Ontario, Canada N5V 4T3  
Tel: 519 457 1470  
Fax: 519 457 3997  
Email: gijzenm@agr.ca.

FEATURES

source  
1. .639  
Location/Qualifiers  
/organism="Glycine max"  
/mol\_type="mRNA"  
/cultivar="Harosoy 63"  
/db\_xref="taxon:3847"  
/tissue\_type="Seed coats"  
/lab\_host="E. coli strain XL0LR"  
/clone\_lib="GmaxSC"  
/note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; This  
cDNA library was constructed from polyA+ enriched mRNA  
from green seed coats in mid to late developmental  
stage, average fresh weight 250 mg per seed. Traces of  
pod and embryo tissue also present. Complementary DNA was  
synthesized from mRNA using an XhoI-poly(dT)  
linker-primer. EcoRI adapters and the products were digested  
blunt-ended cDNA fragments and the products were digested  
with XhoI for directional cloning into lambda ZAP Express  
vector. This lambda library was amplified once using E.  
coli host strain XL1 Blue MRF'. Inserts were then  
subcloned by mass excision using ExAssist helper phage for  
conversion into phagemid vector pBK-CMV in E. coli host  
strain XL0LR."

ORIGIN

Query Match 45.9%; Score 468.4; DB 10; Length 639;  
Best Local Similarity 90.5%; Pred. No. 4e-95;  
Matches 523; Conservative 0; Mismatches 51; Indels 4; Gaps 2;  
Qy 437 GGCAACTATGCGCAAACTGCGAGCTAATGACCTTTTCATTTTGCAGTCCATCTAGAAGG 496  
Db 1 GGAAACTATGCAAACTGCGAGCTAATGATCTTTTCATTTTGCAGGCCATCTAGAAGG 60  
Qy 497 ATTCATATGGAAGTGTGTAGCTGAGTGAATTCAGGATGCTCTCAAGACTACGAG 556  
Db 61 ATTCATATGGAAGTGTGTAGCTGAGTGAATTCAGGATGCTCTCAAGACTACGAG 120  
Qy 557 CTTTAAATTCGAGCTAAGGATAAAGAGATTGATGAATTTGTCATTTCAACAGCGTT 616  
Db 121 CTTTAAATTCGAGCTAAGGATAAAGAGATTGATGAATTTGTCATTTCAACAGCGTT 180  
Qy 617 GAAGACAGCTGAGGAGAGAGTTCGAAAGAGGCTGTGCTTTTGGGAGGAGTGAAT 676  
Db 181 GAAGACAGCTGAGGAGAGAGTTCGAAAGAGGCTGTGCTTTTGGGAGGAGTGAAT 240  
Qy 677 CTTAAGAGTGTGATGACAAATGACAAATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAGC 736

```

Db- 241 CTTGACAATGAGCAGATGATGAGAAACCAATGTTTGCATCCATCAGTGGCTTCTAGC 300
Qy 737 TTGTACAAAATTTGGTGTATACCTCTCACCAGGAGGTTTCCAGTTGAGTACCTCTTGGGC 796
Db 301 TTGTACAAAATTTGGTGTATACCTCTTACCAGGAGGTTTCCAGTTGAGTACCTTATGGG 360
Qy 797 CGTCTGAGTGAAGGCAATTAATGATGAGTGAATTTAGAAAGTGAAGTGAATATGA 856
Db 361 CGTCTGAGTGAAGGCAATTAATGATGAGTGAATTTAGAAAGTGAAGTGAATATGA 420
Qy 857 TGTGTGTTTCAATGATTAAGTCTTTAAGTGTATCTTTTATGCTCAACTTCAATGAGCTG 916
Db 421 TGTGTGTTTCAATGATTAAGTCTTTAAGTGTATCTTTTATGCTCAACTTCAATGAGCTG 480
Qy 917 TTGATTCAGAAATATT-ATTGAGTATGATTTGTTACCTTTATTTTCAATGCTTTAT 975
Db 481 TTGATTAAGAAATATTAAATTTGAGCTTTAGATTTGTTTATTTTCAATGCTTTAT 540
Qy 976 T---AGGAAAGGCAATTTATATATCATGATCTTCAAAA 1010
Db 541 TAGAGAAAAAAGGATTTAGATCAATGAACTTCAATA 578

```

```

RESULT 7
AW774625
LOCUS EST333776 KV3 Medicago truncatula cDNA clone pKV3-23M13, mRNA
DEFINITION sequence.
ACCESSION AW774625
VERSION AW774625.1 GI:7718542
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula

```

```

REFERENCE 1 (bases 1 to 648)
AUTHORS VandenBosch,K., Hurt,J., Moore,J., Beremand,P., Peng,H., Ellis,L.,
Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and
Fraser,C.M.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

```

```

JOURNAL Unpublished (1999)
COMMENT Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@cbs.umn.edu
Texas A&M EST name: T258044e
TIGR sequence name: MTEB379TK
More information is available at:
http://chrysis.tamu.edu/medicago
Seq primer: SKmod (CTA GAA CTA gtc gat CC).
Location/Qualifiers
1..648
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pKV3-23M13"
/tissue_type="Seedling roots"
/dev_stages="3 days post-inoculation with Sinorhizobium
meliloti"
/lab_host="E. coli strain XLOLR"
/clone_lib="KV3"

```

```

FEATURES
source
note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised

```

from the recombinant lambda-Zap phage using Ex-assist  
helper phage and propagated in XLOLR cells."

# ORIGIN

```

Query Match 44.9%; Score 457.6; DB 10; Length 648;
Best Local Similarity 81.6%; Pred. No. 1.1e-92; Indels 0; Gaps 0;
Matches 529; Conservative 0; Mismatches 119;
Qy 122 GGTCTCATTTGAGAGAGCCAAAGTTCCCTTAGCAATCTCACACCTATATGATGAAAAGTATGCT 181
Db 1 GGTCTTATTGAGAGATCAAAAGTTTCCCTTTAATCTCAAACATTATGATCAGATTAATCTT 60
Qy 182 CAATCCAGGGTTTGTGGCTCATTTGGTGGAAATTTGTTGTTAAGAAATACAGAGGCCATT 241
Db 61 CAATCCCTGGTTTGTGGCTCATTTGGTGGAAATTTGTTGTTTCCAAATATCTGAGATTTGT 120
Qy 242 CAAGCTAAGGCTCGAAGATACAAAAACCCCTGAAGAAAACGCCCTTCTTCCAGAAAATTTA 301
Db 121 CAAGCTAAGGCTCGAAGATACAAAAACCCCTGAAGAAATCCCTTCTTCCAGAAAATTTA 180
Qy 302 CCACCATCAATTTGTCATCTTACTCTTCAACACAGTTTTTGCATCCTGGTGTGCTTCA 361
Db 181 TCATTGTCACTTGTTCCTATCTTACCCCTTCAAAAAGTTTTTGCATCCTGGAGCGCTTGG 240
Qy 362 ATTAACATAAACAAGTCCATCTGGAATATGTTTCAAAAGATTTACTTCCATTGCTTGTCT 421
Db 241 ATTAACATAAACAAGTCCATCTGGAATATGTTTCAAAAGATTTACTTCCATTGCTTGTCT 300
Qy 422 ACTTCGGGTGATGATGCGCAACTATGCGCAACTGACGCTTAATGACCTTTTCAATTTGCG 481
Db 301 GCTTCGGGTGATGATGCGCAACTATGCGCAACTGACGCTTAATGACCTTTTCAATTTGCG 360
Qy 482 TCCATCTCTAGAGAGGATTCACATATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCT 541
Db 361 GCATTTCTAAGAGGTTTCAATATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCT 420
Qy 542 CTTCAAGACTACGAGCCTTTAATTCGAGTAAAGGATAAAGAGGATTTGATGAAATTTGTTG 601
Db 421 CTTCAAGACTACGAGCCTTTAATTCGAGTAAAGGATAAAGAGGATTTGATGAAATTTGTTG 480
Qy 602 ACATTTACAGCGTTGAGAGAGCGGTGAGAGAGAGTTGAAAAGAGGCTGTGTGTTT 661
Db 481 ACATTTACAGCGTTGAGAGAGCGGTGAGAGAGAGTTGAAAAGAGGCTGTGTGTTT 540
Qy 662 GGGCAGGAAGTGAATCTTTAACAGTGTATGACATGACATGAAACCGTAAATTTTATCCCA 721
Db 541 GGGCAGGATATAAGCCTTTAACAGCAGTGTATGACATGAAAGTAAACAAAGTTTATCCCA 600
Qy 722 TCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTTCAACCAAG 769
Db 601 TCAGTGGCTTCTAGCTTTGATGAGAAATGGGTAAATACCTCTTACTTAAG 648

```

# RESULT 8

```

BI273174 696 bp mRNA linear EST 18-JUL-2001
NF091B12FL1F1096 Developing flower Medicago truncatula cDNA clone
NF091B12FL 5', mRNA sequence.
BI273174
BI273174.1 GI:14883109
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Medicago truncatula (barrel medic)

```

# REFERENCE

```

1 (bases 1 to 696)
AUTHORS Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLE Expressed sequence tags from the Samuel Roberts Noble Foundation
JOURNAL Medicago truncatula flower library
Unpublished (2001)

```

## COMMENT

Contact: May GD  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel.: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org  
Insert Length: 696 Std Error: 0.00  
Plate: 091 row: B column: 12  
Seq primer: TCACACAGGAACAGCTATGAC.

## FEATURES

source  
1..696  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/db\_xref="taxon:3880"  
/clone="NF091B12P"  
/tissue\_type="Developing flowers"  
/dev\_stage="Developmentally pooled. Contains a mixture of very young, developing, fully-opened flowers and flowers in early transition into pods."  
/clone\_lib="Developing flower"  
/note="Vector: Lambda Zap; cDNA was prepared from polyA+ RNA from very young, developing, fully-opened flowers and flowers transitioning into pods. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

## ORIGIN

Query Match 43.0%; Score 438.2; DB 12; Length 696;  
Best Local Similarity 79.0%; Pred. No. 2.5e-88;  
Matches 543; Conservative 0; Mismatches 142; Indels 2; Gaps 2;  
QY 173 AAGTATGCTCAATTCACAGGTTTGTGGCTCATTTGGTGAATTTGTTTGAAGATACA 232  
Db |||||||  
QY 2 AATTATCTTCAATTCCTCGTTTGTGGCTCATTTGGTGAATTTGTTTCAACAAATAC 61  
Db |||||||  
QY 233 GAGGCCATTCAGTGAAGGTGGAAGATACAAAACCTCGAAGAAAGCGCTTCTCCCA 292  
Db |||||||  
QY 62 GAGATTGTTCAAGTGAAGGTGGAAGATACAAAACCTCGAAGAAAGCGCTTCTCCCA 121  
Db |||||||  
QY 293 GAAATTTACACCATCAATTGTGCCATCTACTCTTCAACAGTTTGTGATCTCGGT 352  
Db |||||||  
QY 122 GAAATTTATCATGTGTCACTGTGTCCATCTTACCCCTTCACAAAGTTTGTGCAATCTGGA 181  
Db |||||||  
QY 353 GCTGCTTCAATTAACATAACAGTCCATCTGGAATGATTTTCAAGAGTTACTTCCA 412  
Db |||||||  
QY 182 GCCGCTTCGATTAACATAACAGTCCCTTATGGAATGATCTTATGACCTGCTTCCA 241  
Db |||||||  
QY 413 TTGCTTGCTACTTCGGGTGATGTGGCAACTATCGCAAACTCGAGCTAATGACCTTTCA 472  
Db |||||||  
QY 242 TTGTTGCTGCTTCGGGTGATGTGGCAACTATGCAAACTCGAGCTACGACCTTCA 301  
Db |||||||  
QY 473 TTATTGAGTCCATCTCTAGAGGATTCATATGGAAGTTTGTAGCTGAGGTGAATTC 532  
Db |||||||  
QY 302 TTGTTGAGGCGCAATTCCTAAAGAGTTTCATTATGGAAGTTTGTAGCTGAGGTGAATTC 361  
Db |||||||  
QY 533 AGGATGCTCTCAAGACTACGAGCTTAAATTCGAGCTAAGGATAAGAGGATTTGATG 592  
Db |||||||  
QY 362 AGGAAATCTCTCAAGACTATGACCTTAAATTCGCTTAAAGACAAAGAGGTTGATG 421  
Db |||||||  
QY 593 AAATTTGTGACATTTACAAGCGTTTGAAGAGACGCTGAGGAGAGAGTTGAAAAGAGGCT 652  
Db |||||||  
QY 422 AAATTTGTGACATTTAAGAGTGTGAAGAGATGTTAAAGAAAGAGGTTGAAGAAAGGCA 481  
Db |||||||  
QY 653 GTGCTGTTTGGCAGGAGTGAATCTTAAAGTGTGACATGACATGACATGAACCCGTAA 712  
Db |||||||  
QY 482 ACATTTGTTGGCAGGATTAAGCCCTTAACAGCAGTGTGACAGTAAAGGTAACAAAG 541  
Db |||||||

QY 713 TTTGATCCATCAGTGGCTTCTAGCTTGATACAAAATTTGGTGATACCTCTCACCAGGAG 772  
Db |||||||  
QY 542 TTTGATCCATCAGTGGCTTCTAGCTTGATACAAAATTTGGTGATACCTCTCACCAGGAG 801  
Db |||||||  
QY 773 GTTC-AGGTTGAGTACCT-CTTGCCCGCTCTAGACTGAAGGCAATTCACATGAGTTGAA 830  
Db |||||||  
QY 602 GGTCAAGGTTGAGTACCTTACTGCCCCCGTTGGATTGAACCAATTTATTTCATGTA 661  
Db |||||||  
QY 831 TTTAGAGAAATGGAAGATGATATGAT 857  
Db |||||||  
QY 662 TTTCAAGAAAGACATATAATATCAT 688  
Db |||||||

## RESULT 9

## AW775842

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## MEDICAGO TRUNCATULA (BARREL MEDIC)

## MEDICAGO TRUNCATULA

## EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; CORE EUDICOTS; ROSIDS; EUROSIDIS I; FABALES; FABACEAE; PAPILIONOIDEAE; TRIFOLIAE; MEDICAGO.

## 1 (bases 1 to 654)

## FEDOROVA, M., PIERSON, B.L., SAMAC, D.A., VANCE, C.P., GANTT, G.S., PENG, H., ELLIS, L., TOWN, C.D., BOWMAN, C.L., CRAVEN, M.B., HANSEN, T.S., HOLC, I.E. and FRASER, C.M.

## ESTs from leaves of Medicago truncatula after inoculation with Colletotrichum trifolii

## Unpublished (2000)

## Contact: Deborah A. Samac

## Department of Plant Pathology

## University of Minnesota

## 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA

## Tel: 612 625 1243

## Fax: 651 649 5058

## Email: debbys@pucini.crl.umn.edu

## Minnesota sequence name: M259175e

## TIGR sequence name: MTEA128TK

## More information is available at:

## http://chryslie.tamu.edu/medicago

## Seq primer: SKmod (CTA GAA CTA Gtg gAT CC).

## Location/Qualifiers

## 1..654

## /organism="Medicago truncatula"

## /mol\_type="mRNA"

## /cultivar="genotype A17"

## /db\_xref="taxon:3880"

## /clone="pDSIL-3E8"

## /tissue\_type="leaves infected with Colletotrichum trifolii"

## /dev\_stage="cotyledons and primary leaves harvested 5 and 9 days after inoculation with Colletotrichum trifolii"

## /lab\_host="E. coli strain XL0LR"

## /clone\_lib="DSIL"

## /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA from cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii. The cDNA was directionally ligated into the Uni-Zap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in XL0LR cells. Note: EST may be of fungal origin."

## ORIGIN

Query Match 42.3%; Score 431.2; DB 10; Length 654;



```

ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 422)
Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 402.
Location/Qualifiers
1 .422
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl032-2453"
/tissue_type="unknown"
/lab_host="unknown"
/clone_lib="Unknown Library Type"
/notes="Vector: unknown; Site 1: unknown; Site 2: unknown;
This library is assigned to clones that are of uncertain
origin due to an error during re-array, prepping, and/or
analysis. Samples assigned this library are still genuine
Glycine Max cDNA sequence, but we are unsure of its source
and identification. Due to the clone id being
incorrectly assigned, THESE CLONES ARE NOT AVAILABLE TO
ORDER."

FEATURES
source
Query Match 40.7%; Score 415.6; DB 12; Length 422;
Best Local Similarity 99.1%; Pred. No. 3.3e-83;
Matches 418; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 357 CTTCAATTACATAAACAAGTCCATCTGGAAGATGATTTCGAAGAGTTACTTCCATTGC 416
DB 1 CTTCAATTACATAAACAAGTCCATCTGGAAGATGATTTCGAAGAGTTACTTCCATTGC 60

QY 417 TTGCTACTTCGGGTGATGTCGCAACTATGCGCAAACTGCAGCTAATGACCTTTTCATTAT 476
DB 61 TTGCTACTTCGGGTGATGTCGCAACTATGCGCAAACTGCAGCTAATGACCTTTTCATTAT 120

QY 477 TGCAGTCCATCTCAGAGGATTCATCTGGAAGTTTGTAGCTAGGTGAATTCAGG 536
DB 121 TGCAGTCCATCTCAGAGGATTCATCTGGAAGTTTGTAGCTAGGTGAATTCAGG 180

QY 537 ATGCTCTCTCAAGACTACGACCTTTTAATTCGAGCTAAGGATAAGAGATTGATGAAT 596
DB 181 ATGCTCTCTCAAGACTACGACCTTTTAATTCGAGCTAAGGATAAGAGATTGATGAAT 240

QY 597 TGTGTGACATTTACAAGCGTTGAAGACCGTGTAGGAAGAGAGTTGAAAGAGCGGTGG 656
DB 241 TGTGTGACATTTACAAGCGTTGAAGACCGTGTAGGAAGAGAGTTGAAAGAGCGGTGG 300

QY 657 TGTTCGGCAGGAAGTGAATCTTTAAGAGTGATGACATGACATGACATGACATGATG 716
DB 301 TGTTCGGCAGGAAGTGAATCTTTAAGAGTGATGACATGACATGACATGACATGATG 360

QY 717 ATCCATCAGTGGCTTCTAGCTTTGTACAAAATTGGGTGATACCTCTCACCAGGAGTTC 776

```

---

```

Db 361 ATCCATCAGTGGCTTCTAGCTTTGTACAAAATTGGGTGATACCTCTCACCAGGAGTTC 420
QY 777 AG 778
Db 421 AG 422

RESULT 12
BM0933460
LOCUS
DEFINITION
560 bp mRNA linear EST 30-NOV-2001
saj09c02.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl065-9724 5' similar to TR:Q9S7H4 Q9S7H4 CHORISMATE MUTASE
; mRNA sequence.
ACCESSION
BM0933460
VERSION
BM0933460.1 GI:17022426
KEYWORDS
Glycine max (soybean)
SOURCE
Glycine max
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 560)
Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 421.
Location/Qualifiers
1 .560
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-9724"
/tissue_type="germinating shoots"
/lab_host="DH10B"
/clone_lib="Gm-cl065"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
germinating shoots of the cultivar Williams. The seeds
were allowed to germinate for 24 hours prior to being
cold stressed for 2 days at 4C. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy
Shoemaker."

ORIGIN
Query Match 39.6%; Score 404; DB 12; Length 560;
Best Local Similarity 87.0%; Pred. No. 1.3e-80;
Matches 487; Conservative 0; Mismatches 0; Indels 73; Gaps 1;

QY 384 GGAAATGTATTTTCAAAGAGTTACTTCCATTGCTTGTACTTGGGTGATGATGGCAACT 443

```





Email: udvardi@pimp-golm.mpg.de

Seq primer: T7

High quality sequence stop: 605.

## FEATURES

Location/Qualifiers

source

1..605  
/organism="Lotus corniculatus var. japonicus"  
/mol\_type="mRNA"  
/cultivar="Gifu (B-129)"  
/db\_xref="taxon:34305"  
/dev\_stage="5 and 7 week-old plants"  
/clone\_lib="Lotus japonicus nodule library 5 and 7 week-old"  
/note="Organ: Nodule; Vector: pSPORT1; Site 1: SalI; Site 2: NotI; The library was prepared using mRNA extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium strain R7A."

## ORIGIN

Query Match 38.1%; Score 388.8; DB 12; Length 605;  
Best Local Similarity 84.2%; Pred. No. 3.4e-77;  
Matches 438; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 48 CTGATTCGGGATGTGTACACCTAGCTCTCTGAGAGAGGATTGGTTAGGCAAGG 107  
Db 86 CAGATCGAATGATGTGTTTCCTCGAGTCTGTGAGGGAGGATCTGGTTAAGCAAGAAG 145  
QY 108 ATACCATCAATTATGTCTCATTTGAGAGAGCCAAAGTTCCTAGCAATTTCTCACACCTATG 167  
Db 146 ATACCATGTTTGTGTTCTCATTTGAGAGGGCTAGTTCCCAATGATTTATCACACCTATG 205  
QY 168 ATGAAAAGTAGTCAAAATCCAGGGTTTGTGGCTCATTTGGTGGAAATTTGTTTAAAG 227  
Db 206 ATGAAACTACTGGAAATCCCGGGTTTGTGGCTCATTTGGTGGAAATTTGTTTCAA 265  
QY 228 ATACAGAGGCAATCAAGCTTAAGCTGGAAGATACAAAACCCCTGAAGAAAAGCCCTTCT 287  
Db 266 ACATGAGCCATCAAGCTATGCTTGGAGATGACAAACCCCTGAAGAAAATGCTTCT 325  
QY 288 TCCGAGAAATTTACCAACATCAATTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 347  
Db 326 TCCGAGAAATTTACCAACATCAATTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 385  
QY 348 CTGCTGCTGCTCAATTAACATAAACAGTCCATCTGGAATGATTTTCAAGAGTTAC 407  
Db 386 CTGAGCTGCTTCCATTAACATAAACAGTCCATCTGGAATGATTTTCAAGAGTTAC 445  
QY 408 TTCCATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 467  
Db 446 TTCCATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 505  
QY 468 TTTCATTTGCAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 527  
Db 506 TTTCATTTGCAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 565  
QY 528 AATTCAGGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 567  
Db 566 AATTCAGGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 605

## RESULT 15

CA800612

LOCUS

DEFINITION

726 bp mRNA linear EST 05-DEC-2002  
sau19a05.y1 Gm-cl062 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl062-8481 5' similar to TR:Q957H4 Q957H4 CHORISMATE MUTASE ;  
mRNA sequence.

ACCESSION

CA800612

VERSION

CA800612.1

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

1 (bases 1 to 726)

## REFERENCE

## AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Corryell,V.,  
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
McCann,R., Waterston,R. and Wilson,R.

## TITLE

## JOURNAL

## COMMENT

Public Soybean EST Project  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact: ccu@resgen.com web site:  
www.resgen.com

Seq primer: -40RP from Gibco  
High quality sequence stop: 409.  
Location/Qualifiers  
1..726

/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-cl062-8481"  
/tissue\_type="stem tissue of greenhouse grown plants"  
/dev\_stage="1 month old"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl062"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from stem tissue of 1 month old greenhouse grown plants  
for the cultivar Raiden. Complementary DNA was  
synthesized from mRNA using a primer consisting of a  
poly(dT) sequence with a XhoI restriction site. EcoRI  
adapters were ligated to the blunt-ended cDNA fragments  
followed by XhoI digestion. The cDNA fragments were  
directionally cloned into the EcoRI-XhoI restriction site  
of the pBluescript vector. The ligated cDNA fragments were  
transformed into DH10B host cells (GibcoBRL). This library  
was constructed in the laboratory of Dr. Randy Shoemaker."

## FEATURES

source

## ORIGIN

Query Match 38.1%; Score 388.8; DB 14; Length 726;  
Best Local Similarity 98.2%; Pred. No. 3.4e-77;  
Matches 393; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 241 TCAAGCTAAGGCTGGAAGATACAAAACCCCTGAAGAAAAGCCCTTCTTCCAGAAAATTT 300  
Db 327 TGATGCACTTGCTGGAAGATACAAAACCCCTGAAGAAAAGCCCTTCTTCCAGAAAATTT 386  
QY 301 ACCACCATCAATTTGCGCATCTTACTCTCTTCAAAACAGTTTTTGCATCCTGCTGCTTC 360  
Db 387 ACCACCATCAATTTGCGCATCTTACTCTCTTCAAAACAGTTTTTGCATCCTGCTGCTTC 446  
QY 361 AATTAAACATAAACAAGTCCATCTGAAAATGTATTTCAAAGAGTTACTTCCATGCTTGC 420  
Db 447 AATTAAACATAAACAAGTCCATCTGAAAATGTATTTCAAAGAGTTACTTCCATGCTTGC 506  
QY 421 TACTTCGGGTGATGATGGCAACTATGCGAAACTGCAAGCTTAATGACCTTTTCAATTATGCA 480  
Db 507 TACTTCGGGTGATGATGGCAACTATGCGAAACTGCAAGCTTAATGACCTTTTCAATTATGCA 566  
QY 481 GTCCATCTCTAGAGAGGATTCACATATGGAAGTTTGTAGCTGAGGTGAATTCAGGATGC 540  
Db 567 GTCCATCTCTAGAGAGGATTCACATATGGAAGTTTGTAGCTGAGGTGAATTCAGGATGC 626  
QY 541 TCCTCAAGACTACGAGCCTTTAATTCAGCTAAGGATAAAGAGGATTGTAATTTGTT 600

us-10-624-061-15.rst

Mon Sep 27 07:44:33 2004

Db 627 TCCTCAAGACTACGAGCCTTTAATTCGAGCTAAGATTAAGAGGATTTGATGAATTGTT 686  
Qy 601 GACATTTACAAGCGTTGAAGAGACGCGTCAGGAAGAGATT 640  
Db 687 GACATTTACAAGCGTTGAAGAGACGCGGAGGAAGAGATT 726

Search completed: September 25, 2004, 01:24:49  
Job time : 3099 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2004, 01:24:55 ; Search time 74 Seconds  
(without alignments)  
996.553 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339  
Sequence: 1 MAXAAEQSPDSGNVYTLASV.....NWVPLTKVQVEYLLRLD 261

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1339	100.0	261	6	ABU08094 Soybean c
2	818	61.1	265	5	AAM49654 A. thalia
3	818	61.1	265	5	ABB3424 Herbicida
4	818	61.1	285	6	ABU08098
5	818	61.1	330	5	AAM49655 Construct
6	698	52.1	334	5	ABB32474 Herbicida
7	696	52.0	334	6	ABU08097
8	681.5	50.9	257	6	ABU08093 Rice chor
9	673	50.3	132	6	ABU08089
10	670.5	50.1	311	6	ABU08092
11	669.5	50.0	316	5	ABB31535
12	602	45.0	207	6	ABU08095
13	599.5	44.8	319	3	AAG06003 Arabidops
14	599.5	44.8	325	3	AAG06002
15	599.5	44.8	347	3	AAG06001 Arabidops
16	489	36.5	154	6	ABU08090
17	424.5	31.7	263	6	AAO19566 Protein e
18	423	31.6	280	3	AAO19566 M sterili
19	147	11.0	93	6	ABU08087
20	131	9.8	62	6	ABU08088
21	105	7.8	360	6	ABU00704
22	103	7.7	781	4	AAU35676 Haemophil
23	103	7.7	781	6	ABU30650 Protein e
24	96.5	7.2	655	4	ABG25473 Novel hum
25	96	7.2	459	4	AAG2076 S. epider

26	96	7.2	459	4	AAG2730 S. epider
27	96	7.2	464	5	ABP39856 Staphyloc
28	95.5	7.1	594	6	ABM73376 Staphyloc
29	94	7.0	1037	4	ABG04999 Novel hum
30	93.5	7.0	739	6	ABU25591 Protein e
31	92.5	6.9	537	4	ABBS7939 Prosophil
32	92	6.9	605	6	ABBS7939 Protein s
33	91.5	6.8	433	6	ABU58233
34	91.5	6.8	649	6	ADA20823 Rice stre
35	91.5	6.8	676	6	ABU25397 Protein e
36	90.5	6.8	554	2	AAM69223 HAP4 prot
37	90	6.7	693	6	ABU44329 Protein e
38	89.5	6.7	590	7	ADC95055 E. faeciu
39	88.5	6.6	680	6	ADB23150 Environme
40	88.5	6.6	856	3	AA341660
41	88.5	6.6	1173	4	AA979111 Human ORF
42	88.5	6.6	1948	5	ABP69390 Human G-p
43	88.5	6.6	1958	7	ADE14371 Human int
44	88	6.6	1072	2	AAV29460 African h
45	87.5	6.5	596	6	ABU24103 Protein e

## ALIGNMENTS

RESULT 1  
ABU08094

ID ABU08094 standard; protein; 261 AA.

XX AC ABU08094;  
XX AC ABU08094;

DT 22-MAY-2003 (first entry)  
XX XX

DE DE Soybean chorismate mutase protein.  
XX XX

KW Soybean; plant; enzyme; chorismate mutase; aromatic amino acid; diet;  
KW biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;  
KW phenylalanine; prephenate dehydrogenase; herbicide; food crop.  
XX OS Glycine max.  
XX XX

PN US2002184658-A1.  
XX XX

PD 05-DEC-2002.  
XX XX

PF 03-DEC-1999; 99US-00454279.  
XX XX

PR 04-DEC-1998; 98US-0110845P.  
XX XX

PA (FALC/) FALCO S. C.  
PA (FAMO/) FAMODU O. O.

XX (LEEJ/) LEE J.  
XX XX

PI FALCO SC, Famodu OO, Lee J;  
XX XX

DR WPI: 2003-328651/31.  
XX XX

DR N-PSDB; ABX93058.  
XX XX

PT New tyrosine biosynthetic enzyme, chorismate mutase, proteins and  
PT nucleic acids, useful for facilitating design and/or identifying  
PT inhibitors of those enzymes that may be used as herbicides and for  
producing antibodies.

PS Claim 11; Fig 1; 32pp; English.  
XX XX

CC The invention discloses isolated polynucleotides encoding chorismate  
CC mutase polypeptides. Aromatic amino acids must be included in the diets  
CC of animals. In the aromatic amino acid biosynthetic pathway chorismate is  
CC converted to anthranilate during tryptophan biosynthesis and is converted  
CC to prephenate, the branch point for tyrosine and phenylalanine  
CC biosynthesis. Chorismate mutase catalyses the conversion of chorismate to  
CC prephenate. Also disclosed are methods for selecting an isolated  
CC polynucleotide that affects the level of expression of a tyrosine

CC biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
CC fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
CC evaluating compounds for their ability to inhibit the activity of a  
CC tyrosine biosynthetic enzyme. The polypeptides can be used produce  
CC antibodies. Chorismate mutase and prephenate dehydrogenase are good  
CC targets for herbicides that will not affect animals, and overexpression  
CC of these enzymes may be used to increase the content of aromatic amino  
CC acid in food crops. The polypeptides may also be used to design and/or  
CC identify inhibitors of those enzymes that may be used as herbicides. The  
CC nucleic acids may be used to create transgenic plants, as probes for the  
CC genetic and physical mapping of the genes and as markers for traits  
CC linked to those genes. The sequence presented is a soybean chorismate  
CC mutase protein  
XX  
XX  
SQ Sequence 261 AA;

Query Match 100.0%; Score 1339; DB 6; Length 261;  
Best Local Similarity 100.0%; Pred. No. 2.2e-133;  
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAKAAEQSDSGNVYTLASVREDLVROEDTIYGLIERAKPSPNSHTYDEKYAQIQGFCG 60  
DB 1 MAKAAEQSDSGNVYTLASVREDLVROEDTIYGLIERAKPSPNSHTYDEKYAQIQGFCG 60  
QY 61 SLVEFVVKNTETAIQAKAGRYKNPEENAFPPENLPSPISVPSYFQFLHPGAASININKSI 120  
DB 61 SLVEFVVKNTETAIQAKAGRYKNPEENAFPPENLPSPISVPSYFQFLHPGAASININKSI 120  
QY 121 WKMYFKELLPLATSGDDGNYAQTAANDLSLQSRRIHYGKFAEVKFRDAPQDYEP 180  
DB 121 WKMYFKELLPLATSGDDGNYAQTAANDLSLQSRRIHYGKFAEVKFRDAPQDYEP 180  
QY 181 IRAKDKKGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPSPVASSLY 240  
DB 181 IRAKDKKGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPSPVASSLY 240  
QY 241 KKWVPLTKREVQVEYLLRLD 261  
DB 241 KKWVPLTKREVQVEYLLRLD 261

RESULT 2  
ID AAM49654  
XX AAM49654 standard; protein; 265 AA.  
AC AAM49654;  
DT 17-MAY-2002 (first entry)  
XX A. thaliana cytosolic chorismate mutase-2 protein.  
DE Shikimate pathway; chorismate mutase; prephenate dehydrogenase;  
XX vitamin E; vitamin K; ubiquinone; fat-soluble antioxidant; transgenic;  
KW feed; pharmaceutical; cosmetic; aromatic amino acid; salicylic acid;  
KW folic acid; phenylpropanoid; flavanoid; plant; abiotic stress resistance;  
KW tocopherol; tocotrienol.  
XX Arabidopsis thaliana.  
XX OS Arabidopsis thaliana.  
XX PN WO200200901-A1.  
XX PD 03-JAN-2002.  
XX PF 28-JUN-2001; 2001WO-EP007391.  
XX PP 29-JUN-2000; 2000DE-01030647.  
XX PR 21-DEC-2000; 2000DE-01064454.  
XX PA (SUNG-) SUNGENE GMBH & CO KGAA.  
XX FI Badur R, Geiger M, Kunze I, Sommer S;  
XX WPI; 2002-164442/21.  
XX DR

DR N-PSDB; ABA99613.  
XX Preparing fine chemicals, particularly Vitamins E and K, useful as  
PT antioxidants e.g. in foods or medicine, by growing organisms with altered  
PT shikimate biosynthesis pathway.  
XX  
PS Claim 9; Page 60-61; 74pp; German.  
XX This invention describes a novel method for the preparation of fine  
CC chemicals by culturing organisms in which the shikimate pathway has been  
CC altered relative to the wild type. The method involves a construct  
CC containing a plasmid trans peptide, a chorismate mutase and/or  
CC prephenate dehydrogenase, linked to at least one regulatory sequence for  
CC transcription or translation in plants. The method is used to produce  
CC chemicals, particularly Vitamins E or K and/or ubiquinone, useful e.g. as  
CC fat-soluble antioxidants (in feeds, pharmaceuticals and cosmetics), but  
CC also aromatic amino acids, salicylic or folic acid derivatives,  
CC phenylpropanoids, flavanoids etc., especially in transgenic plants. Also  
CC plants with increased Vitamin E contents have improved resistance to  
CC abiotic stress, e.g. frost and drought. Transgenic plants with an altered  
CC shikimate pathway are useful as foods, fodder and in preparation of  
CC processed foodstuffs. Transgenic plants with modified shikimate pathways  
CC have increased content of desired chemicals, particularly tocopherols  
CC and/or tocotrienols. This sequence represents the Arabidopsis thaliana  
CC cytosolic chorismate mutase-2 protein used in the method of the invention  
XX  
XX SQ Sequence 265 AA;

Query Match 61.1%; Score 818; DB 5; Length 265;  
Best Local Similarity 61.1%; Pred. No. 5.3e-78;  
Matches 162; Conservative 31; Mismatches 68; Indels 4; Gaps 2;  
QY 1 MAKAAEQSDSG--NVYTLASVREDLVROEDTIYGLIERAKPSPNSHTYDEKYAQIQG 58  
DB 1 MARVFSDSGSGSNVLSLRLRESLROEDTIVFSLIERAKPPLNSPAPFESRCLDSGS 60  
QY 59 CGSLVEFVVKNTETAIQAKAGRYKNPEENAFPPENLPSPISVPSYFQFLHPGAASININK 118  
DB 61 FSSLTFFVRETIQAKGRVEYEPENPFLENIPHSVFTHKYPSALHPKALSININK 120  
QY 119 SIWKMYFKELLPLATSGDDGNYAQTAANDLSLQSRRIHYGKFAEVKFRDAPQDY 178  
DB 121 QIWDIVFKELLPLFVKPGDDGNYPTAASDLACLQALSRIHYGKFAEVKFRDAPQDY 180  
QY 179 PLIRAKDKKGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKF--DPSVA 236  
DB 181 PAIRAQDREALMKLLTFEKEVSENVKVEKQKKAETFGQEVKNSGYGDESKKVKYDPLLA 240  
QY 237 SSLLYKQWVPLTKREVQVEYLLRLD 261  
DB 241 SRIYGEWLIPLTKREVQVEYLLRLD 265

RESULT 3  
ID ABB93424  
XX ABB93424 standard; protein; 265 AA.  
AC ABB93424;  
XX 31-MAY-2002 (first entry)  
XX Herbicidally active polypeptide SEQ ID NO 2635.  
XX DE Herbicidal; plant; agriculture; herbicide.  
XX KW Arabidopsis thaliana.  
XX OS Arabidopsis thaliana.  
XX PN WO200210210-A2.  
XX PD 07-FEB-2002.  
XX PF 28-AUG-2001; 2001WO-EP009892.  
XX PP  
XX DR



QY	179	PLIRAKDKEGIMKLLTTSVEETVRKRVEKKAUVFGQEVNLSNDNDNENRKF--DPSVA	233
Db	246	PAIRAQDREALKLLITEKVEVMVKVKQKAEFTFOEVKFNKSGYGESEKKYKVDPLLA	305
QY	237	SSLYKNWVPIPLTKVEQVVEYLLRLD	261
Db	306	SRIGEWLPIPLTKLVEVEYLLRLD	330
RESULT 6			
ABB92474	ABB92474 standard; protein; 334 AA.		
XX	AC	ABB92474;	
XX	DT	31-MAY-2002 (first entry)	
XX	DE	Herbicidally active polypeptide SEQ ID NO 1685.	
XX	KW	Herbicidal; plant; agriculture; herbicide.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO200210210-A2.	
XX	PN	07-FEB-2002.	
XX	PD	28-AUG-2001; 2001WO-EP009892.	
XX	PF	28-AUG-2001; 2001WO-EP009892.	
XX	PR	(FARB ) BAYER AG.	
XX	PA	Tietjen K, Weidler M;	
XX	PI	WPI; 2002-269010/31.	
XX	DR	Identifying plant target proteins for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms.	
XX	PS	Claim 5; SEQ ID NO 1685; 261pp + Sequence Listing; English.	
XX	CC	The invention relates to identifying target proteins (ABB90790-ABB94016) for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides	
XX	XX	Sequence 334 AA;	
XX	XX	Query Match 52.1%; Score 698; DB 5; Length 334;	
XX	XX	Best Local Similarity 50.4%; Pred. No. 4.2e-65;	
XX	XX	Matches 131; Conservative 46; Mismatches 75; Indels 8; Gaps 1	
QY	10	DSGNVYTLASVREDIVROEDTIIYGLTERAKFNSHTYDEKVAQIQGCGSLVEFVVKH.69	69
Db	75	DSESLTLEGIRNSLIQEDSIIIGLLERAKYCNADTYPTAFDMDGFGNSLVEVWKG	134
QY	70	TEAIOAKAGRYKNPEENAPFENLPPSIVPSYFKOFLHFGAASININKSIWQYFKELL	129
Db	135	TEKLHAKVGRFSPDEHPFPDDLPPEMLPPLQPVKVLHFAADSININKKIWNMYPRDIV	194
QY	130	PLLATSGDGNVQATANDLSLOSISRRIHYGKFAEVKFRDAPODYPLTRAKDKEGI	189
Db	195	PLRVKKGDDGNYGSTAVCDALICLOLSKR.IHYGKFAEVAKFOASPEAYESAIAQDKDAL	254
QY	190	MKLLTFTSVETVRKRVEKKAUVFGQEVNLSNDNDNENR-----KFPDPSVASSLYK	241

Db	255	MDMLTFTPTVEDAIKRVEMKTRTYGOEVKVGMEKEEBEENESHVYKISPIVLGDLYG	314
Qy	242	NWVPLTKVQVEYLLRLD	261
Db	315	DWIMPLTKVQVEYLLRLD	334
RESULT 7			
ABU08097			
ID	ABU08097	standard; protein; 334 AA.	
XX	ABU08097;		
XX	22-MAY-2003	(first entry)	
DT	22-MAY-2003	(first entry)	
XX	A. thaliana chorismate mutase protein #1.		
DE	Plant; enzyme; chorismate mutase; aromatic amino acid; diet;		
KW	biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;		
KW	phenylalanine; prephenate dehydrogenase; herbicide; food crop.		
XX	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		
XX	US2002184658-A1.		
PN	US2002184658-A1.		
XX	05-DEC-2002.		
PD	05-DEC-2002.		
XX	03-DEC-1999;	99US-00454279.	
PF	03-DEC-1999;	99US-00454279.	
XX	04-DEC-1998;	98US-0110845P.	
PR	04-DEC-1998;	98US-0110845P.	
XX	(FALC/) FALCO S. C.		
PA	(FAMO/) FAMODU O O.		
PA	(LEEU/) LEE J.		
XX	Falco SC, Famodu OO, Lee J;		
PI	Falco SC, Famodu OO, Lee J;		
XX	WPI; 2003-328651/31.		
DR	WPI; 2003-328651/31.		
XX	New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and		
PT	nucleic acids, useful for facilitating design and/or identifying		
PT	inhibitors of those enzymes that may be used as herbicides and for		
PT	producing antibodies.		
XX	Example 3; Fig 1; 32pp; English.		
PS	Example 3; Fig 1; 32pp; English.		
XX	The invention discloses isolated polynucleotides encoding chorismate		
CC	mutase polypeptides. Aromatic amino acids must be included in the diets		
CC	of animals. In the aromatic amino acid biosynthetic pathway chorismate is		
CC	converted to anthranilate during tryptophan biosynthesis and is converted		
CC	to prephenate, the branch point for tyrosine and phenylalanine		
CC	biosynthesis. Chorismate mutase catalyzes the conversion of chorismate to		
CC	prephenate. Also disclosed are methods for selecting an isolated		
CC	polynucleotide that affects the level of expression of a tyrosine		
CC	biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid		
CC	fragment encoding a tyrosine biosynthetic enzyme polypeptide and		
CC	evaluating compounds for their ability to inhibit the activity of a		
CC	tyrosine biosynthetic enzyme. The polypeptides can be used produce		
CC	antibodies. Chorismate mutase and prephenate dehydrogenase are good		
CC	targets for herbicides that will not affect animals, and overexpression		
CC	of these enzymes may be used to increase the content of aromatic amino		
CC	acid in food crops. The polypeptides may also be used to design and/or		
CC	identify inhibitors of those enzymes that may be used as herbicides. The		
CC	nucleic acids may be used to create transgenic plants, as probes for the		
CC	genetic and physical mapping of the genes and as markers for traits		
CC	linked to those genes. The sequence presented is the A. thaliana		
CC	chorismate mutase protein #1		
XX	Sequence 334 AA;		
QQ	Sequence 334 AA;		
Query Match			
Best Local Similarity			
		52.0%;	Score 696; DB 6; Length 334;
		50.4%;	Pred. No. 6.9e-65;

Matches	131;	Conservative	46;	Mismatches	75;	Indels	8;	Gaps	1;
Qy	10	DSGNVYTTASVREDLVROEDTIYGLIERAFKFPNSHTYDEKYAQIQGFCGSLVEFVYKN	69						
Db	75	DESESLTLEGIRNSLIROEDSIIIFGLERAKYCYNADTYDPTAFMDGFGNSLVEYVWVG	134						
Qy	70	TEAIQAKAGRYKNPEENAFPPENPPSVPSYSPKQFLHPOGAAININIKSWKMYFKELL	129						
Db	135	TEKLHAKVGREKSPDEHPFFDDLPPEPMLPQPKVLHFAADSIININKKIWNMYFRDLV	194						
Qy	130	PLLATSGDDGNYAOTAAANDLSLLOSISRRIHYGKFAEVEKPRDAPQDYVEPLIRAKDKEGL	189						
Db	195	PLVYKGGDDGNYGSTAVCDALICQLSKRIHYGKFAEVEKFAQSPAYESAIAKQDRDL	254						
Qy	190	MKLLTFTSVEETVRKRVKKAIVFGQEVNLSNDNDNENR-----KDPSPVASSLIYK	241						
Db	255	MDMLTFTPTVEDAIKRVEMKTRTYGOEVKVGMEKEEBEENESHVYKISPIVLGDLYG	314						
Qy	242	NWVPLTKVQVEYLLRLD	261						
Db	315	DWIMPLTKVQVEYLLRLD	334						
RESULT 8									
ABU08093									
ID	ABU08093	standard; protein; 257 AA.							
XX	ABU08093;								
XX	22-MAY-2003	(first entry)							
DT	22-MAY-2003	(first entry)							
XX	Rice chorismate mutase protein.								
DE	Rice chorismate mutase protein.								
XX	Rice; plant; enzyme; chorismate mutase; aromatic amino acid; diet;								
KW	biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;								
KW	phenylalanine; prephenate dehydrogenase; herbicide; food crop.								
XX	Oryza sativa.								
OS	Oryza sativa.								
XX	US2002184658-A1.								
PN	US2002184658-A1.								
XX	05-DEC-2002.								
PD	05-DEC-2002.								
XX	03-DEC-1999;	99US-00454279.							
PF	03-DEC-1999;	99US-00454279.							
XX	04-DEC-1998;	98US-0110845P.							
PR	04-DEC-1998;	98US-0110845P.							
XX	(FALC/) FALCO S. C.								
PA	(FAMO/) FAMODU O O.								
PA	(LEEU/) LEE J.								
XX	Falco SC, Famodu OO, Lee J;								
PI	Falco SC, Famodu OO, Lee J;								
XX	WPI; 2003-328651/31.								
DR	WPI; 2003-328651/31.								
XX	New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and								
PT	nucleic acids, useful for facilitating design and/or identifying								
PT	inhibitors of those enzymes that may be used as herbicides and for								
PT	producing antibodies.								
XX	Claim 11; Fig 1; 32pp; English.								
PS	Claim 11; Fig 1; 32pp; English.								
XX	The invention discloses isolated polynucleotides encoding chorismate								
CC	mutase polypeptides. Aromatic amino acids must be included in the diets								
CC	of animals. In the aromatic amino acid biosynthetic pathway chorismate is								
CC	converted to anthranilate during tryptophan biosynthesis and is converted								
CC	to prephenate, the branch point for tyrosine and phenylalanine								
CC	biosynthesis. Chorismate mutase catalyzes the conversion of chorismate to								
CC	prephenate. Also disclosed are methods for selecting an isolated								
CC	polynucleotide that affects the level of expression of a tyrosine								
CC	biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid								
CC	fragment encoding a tyrosine biosynthetic enzyme polypeptide and								
CC	evaluating compounds for their ability to inhibit the activity of a								
CC	tyrosine biosynthetic enzyme. The polypeptides can be used produce								
CC	antibodies. Chorismate mutase and prephenate dehydrogenase are good								
CC	targets for herbicides that will not affect animals, and overexpression								
CC	of these enzymes may be used to increase the content of aromatic amino								
CC	acid in food crops. The polypeptides may also be used to design and/or								
CC	identify inhibitors of those enzymes that may be used as herbicides. The								
CC	nucleic acids may be used to create transgenic plants, as probes for the								
CC	genetic and physical mapping of the genes and as markers for traits								
CC	linked to those genes. The sequence presented is the A. thaliana								
CC	chorismate mutase protein #1								
XX	Sequence 334 AA;								
QQ	Sequence 334 AA;								
Query Match									
Best Local Similarity									
		52.0%;	Score 696; DB 6; Length 334;						
		50.4%;	Pred. No. 6.9e-65;						

CC evaluating compounds for their ability to inhibit the activity of a  
 CC tyrosine biosynthetic enzyme. The polypeptides can be used produce  
 CC antibodies. Chorismate mutase and prephenate dehydrogenase are good  
 CC targets for herbicides that will not affect animals, and overexpression  
 CC of these enzymes may be used to increase the content of aromatic amino  
 CC acid in food crops. The polypeptides may also be used to design and/or  
 CC identify inhibitors of those enzymes that may be used as herbicides. The  
 CC nucleic acids may be used to create transgenic plants, as probes for the  
 CC genetic and physical mapping of the genes and as markers for traits  
 CC linked to those genes. The sequence presented is a rice chorismate mutase  
 CC protein

XX Sequence 257 AA;  
 CC  
 CC Query Match 50.9%; Score 681.5; DB 6; Length 257;  
 CC Best Local Similarity 52.1%; Pred. No. 1.6e-63;  
 CC Matches 134; Conservative 44; Mismatches 78; Indels 1; Gaps 1;  
 CC  
 CC QY 6 EQSPSGNVYTLASVREDLVROEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCGLVEF 65  
 CC DB 1 KORIDQSELTLDNRTSLVRQEDSIIFSLERAQFCYNADIYDKNAFVDFGDSLVEF 60  
 CC  
 CC QY 66 VVKNTAEIAQKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGGAASININKSIWKMYF 125  
 CC DB 61 MYRETEKLHQVGRYKSPDEHPFFEDLPEPLLPQYPKVLHPITADSNINKELWKMYF 120  
 CC  
 CC QY 126 KELLPLLATSGDDGNYAQTAAANDLSLOSISRHHYGFVAEVKFRDAPQDYELIRAKD 185  
 CC DB 121 DELLPLRVKSGDNGYSSALCDTICLQALSRIHYGRFVAEAKFQESPEAYMPAIAQD 180  
 CC  
 CC QY 186 KEGLMKLLFTVEETVRKVEKKAIVFGQEVNLSDDNENR-KFPDPSVASSLYKNGV 244  
 CC DB 181 CQQLMELLTYETVERAIEHREKAKIKFGQEVLDGAENGAPPYMKIRPSLVAELYSVRI 240  
 CC  
 CC QY 245 IPLTKEVQVEYLLRLID 261  
 CC DB 241 MPLTKEVQVAYLLRLID 257

RESULT 9  
 ABU08089  
 ID ABU08089 standard; protein; 132 AA.  
 AC ABU08089;  
 DT 22-MAY-2003 (first entry)  
 XX Protein encoded by soybean chorismate mutase expressed sequence tag.  
 DE Soybean; plant; enzyme; chorismate mutase; expressed sequence tag; EST;  
 XX aromatic amino acid; diet; biosynthesis; chorismate; anthranilate;  
 KW tryptophan; prephenate; tyrosine; phenylalanine;  
 KW prephenate dehydrogenase; herbicide; food crop.  
 XX Glycine max.  
 OS  
 XX Key Location/Qualifiers  
 FH FT Misc-difference 130  
 FT FT /note= "Encoded by GGN"  
 XX  
 PN US2002184658-A1.  
 XX  
 XX 05-DEC-2002.  
 XX  
 XX 03-DEC-1999; 99US-00454279.  
 XX  
 XX 04-DEC-1998; 98US-0110845P.  
 XX  
 XX (FALC/) FALCO S C.  
 PA (FAMC/) FAMODU O O.  
 PA (LEEJ/) LEE J.  
 XX  
 XX Falco SC, Famodu OO, Lee J;

XX WPI; 2003-328651/31.  
 DR N-PSDB; ABX93053.  
 XX  
 XX New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
 PT nucleic acids, useful for facilitating design and/or identifying  
 PT inhibitors of those enzymes that may be used as herbicides and for  
 PT producing antibodies.  
 XX  
 XX Claim 11; Page 16-17; 32pp; English.  
 XX  
 CC The invention discloses isolated polynucleotides encoding chorismate  
 CC mutase polypeptides. Aromatic amino acids must be included in the diets  
 CC of animals. In the aromatic amino acid biosynthetic pathway chorismate is  
 CC converted to anthranilate during tryptophan biosynthesis and is converted  
 CC to prephenate, the branch point for tyrosine and phenylalanine  
 CC biosynthesis. Chorismate mutase catalyses the conversion of chorismate to  
 CC prephenate. Also disclosed are methods for selecting an isolated  
 CC polynucleotide that affects the level of expression of a tyrosine  
 CC biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
 CC fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
 CC evaluating compounds for their ability to inhibit the activity of a  
 CC tyrosine biosynthetic enzyme. The polypeptides can be used produce  
 CC antibodies. Chorismate mutase and prephenate dehydrogenase are good  
 CC targets for herbicides that will not affect animals, and overexpression  
 CC of these enzymes may be used to increase the content of aromatic amino  
 CC acid in food crops. The polypeptides may also be used to design and/or  
 CC identify inhibitors of those enzymes that may be used as herbicides. The  
 CC nucleic acids may be used to create transgenic plants, as probes for the  
 CC genetic and physical mapping of the genes and as markers for traits  
 CC linked to those genes. The sequence presented is the protein encoded by a  
 CC soybean chorismate mutase expressed sequence tag (EST)  
 XX  
 XX Sequence 132 AA;

Query Match 50.3%; Score 673; DB 6; Length 132;  
 Best Local Similarity 98.5%; Pred. No. 4.6e-63;  
 Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 10 DSGNVYTLASVREDLVROEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCGLVEFVKN 69  
 DB 1 DSGNVYTLASVREDLVROEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCGLVEFVKN 60  
 QY 70 TEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGGAASININKSIWKMYFKELL 129  
 DB 61 TEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGGAASININKSIWKMYFKELL 120  
 QY 130 PLLATSGDDGNY 141  
 DB 121 PLLATSGDDGNY 132  
 RESULT 10  
 ABU08092  
 ID ABU08092 standard; protein; 331 AA.  
 AC ABU08092;  
 XX  
 XX 22-MAY-2003 (first entry)  
 XX  
 XX Corn chorismate mutase protein.  
 DE  
 XX Corn; plant; enzyme; chorismate mutase; aromatic amino acid; diet;  
 KW biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;  
 KW phenylalanine; prephenate dehydrogenase; herbicide; food crop.  
 XX  
 XX Zea mays.  
 XX  
 XX US2002184658-A1.  
 XX  
 XX 05-DEC-2002.  
 XX  
 XX 03-DEC-1999; 99US-00454279.



```

XX PR 04-DEC-1998; 98US-0110845P.
XX PA (FALC/) FALCO S C.
XX PA (FAMO/) FAMODU O O.
XX PA (LEEU/) LEE J.
XX PI Falco SC, Famodu OO, Lee J;
XX WPI; 2003-328651/31.
XX DR N-PSDB; ABX93056.
XX PT New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and
XX PT nucleic acids, useful for facilitating design and/or identifying
XX PT inhibitors of those enzymes that may be used as herbicides and for
XX PT producing antibodies.
XX PS Claim 11; Fig 1; 32pp; English.
XX CC The invention discloses isolated polynucleotides encoding chorismate
XX CC mutase polypeptides. Aromatic amino acids must be included in the diets
XX CC of animals. In the aromatic amino acid biosynthetic pathway chorismate is
XX CC converted to anthranilate during tryptophan biosynthesis and is converted
XX CC to prephenate, the branch point for tyrosine and phenylalanine
XX CC biosynthesis. Chorismate mutase catalyses the conversion of chorismate to
XX CC prephenate. Also disclosed are methods for selecting an isolated
XX CC polynucleotide that affects the level of expression of a tyrosine
XX CC biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid
XX CC fragment encoding a tyrosine biosynthetic enzyme polypeptide and
XX CC evaluating compounds for their ability to inhibit the activity of a
XX CC tyrosine biosynthetic enzyme. The polypeptides can be used produce
XX CC antibodies. Chorismate mutase and prephenate dehydrogenase are good
XX CC targets for herbicides that will not affect animals, and overexpression
XX CC of these enzymes may be used to increase the content of aromatic amino
XX CC acid in food crops. The polypeptides may also be used to design and/or
XX CC identify inhibitors of those enzymes that may be used as herbicides. The
XX CC nucleic acids may be used to create transgenic plants, as probes for the
XX CC genetic and physical mapping of the genes and as markers for traits
XX CC linked to those genes. The sequence presented is a corn chorismate mutase
XX CC protein
XX SQ Sequence 331 AA;
Query Match 50.1%; Score 670.5; DB 6; Length 331;
Best Local Similarity 51.4%; Pred. No. 3.5e-62;
Matches 130; Conservative 48; Mismatches 74; Indels 1; Gaps 1;
QY 10 DSGNVYTLASVREDLVQRQEDTIYGLIERAKFPNSHTYDEKYAQIQGFGSLVEFVVK 69
DB 79 DRSDILTDLSIRQVLRLEDSIFGLIERAQFCYNADTYDSNAFMDGFGSLVEIVRE 138
QY 70 TEATQAKAGRYKNPEENAFFPENLPPSIVPSYFKQFLHPGAASININKSIWQYFKELL 129
DB 139 TEKLHAQVGRYKSPDEHPFSSKOLPEPLPMQYPRVLPHPIDADSINKEIWKQYFDELL 198
QY 130 PLLATSGDDGNYAQTAANDLSLQSISSRIHYGKFAEVKFRDAPQDYELIRAKDKEGL 189
DB 199 PRLVKGSGDGNAGSALCDTTCIQASRRRIHYGKFAEVKFRDAPQDYELIRAKDKEGL 258
QY 190 MKLITFTSVETVAKRVEKKAIVFGQEVNLSNDNEN-RKFDPSVASSLYKNWVPLT 248
DB 259 MLLITYETVERAIEHVEAKRIFGQEVNIGAKDNGSPVYKIRPSLVAELYSYRIMPLT 318
QY 249 KEVOVEYLLRLD 261
DB 319 KEVEVAYLLKRLD 331
RESULT 11
ABB91515
ID ABB91515 standard; protein; 316 AA.
XX AC ABB91515;

```

```

XX DT 31-MAY-2002 (first entry)
XX DE Herbicidally active polypeptide SEQ ID NO 726.
XX KW Herbicidal; plant; agriculture; herbicide.
XX OS Arabidopsis thaliana.
XX FN WO200210210-A2.
XX PD 07-FEB-2002.
XX PF 28-AUG-2001; 2001WO-EP009892.
XX PR 28-AUG-2001; 2001WO-EP009892.
XX PA (FARB ) BAYER AG.
XX PI Tietjen K, Weidler M;
XX WPI; 2002-269010/31.
XX PT Identifying plant target proteins for herbicidally active compounds,
XX PT comprising aligning and comparing nucleic acid or amino acid sequences
XX PT from plant with nucleic acid or amino acid sequences from non-plant
XX PT organisms.
XX PS Claim 5; SEQ ID NO 726; 261pp + Sequence Listing; English.
XX CC The invention relates to identifying target proteins (ABB90790-ABB94016)
XX CC for herbicidally active compounds, comprising aligning and comparing
XX CC nucleic acid or amino acid sequences from plant with nucleic acid or
XX CC amino acid sequences from non-plant organisms using suitable search
XX CC parameters, where plant sequences having an E-value greater by a factor
XX CC of 3 than the E-value of most similar non-plant sequences are selected.
XX CC The polypeptides or nucleic acids encoding them are useful for
XX CC identifying modulators. The identified modulators are useful as
XX CC herbicides
XX SQ Sequence 316 AA;
Query Match 50.0%; Score 669.5; DB 5; Length 316;
Best Local Similarity 51.4%; Pred. No. 4.1e-62;
Matches 130; Conservative 44; Mismatches 78; Indels 1; Gaps 1;
QY 10 DSGNVYTLASVREDLVQRQEDTIYGLIERAKFPNSHTYDEKYAQIQGFGSLVEFVVK 69
DB 64 DESEYKLKESIRHSLIROEDSIIFNLERAQRYNADTYDEDAFTMEGFGSLVEFVWRE 123
QY 70 TEATQAKAGRYKNPEENAFFPENLPPSIVPSYFKQFLHPGAASININKSIWQYFKELL 129
DB 124 TEKLHAQVGRYKSPDEHPFSSKOLPEPLPMQYPRVLPHPIDADSINKEIWKQYFDELL 183
QY 130 PLLATSGDDGNYAQTAANDLSLQSISSRIHYGKFAEVKFRDAPQDYELIRAKDKEGL 189
DB 184 PRLVKGSGDGNAGSALCDTTCIQASRRRIHYGKFAEVKFRDAPQDYELIRAKDKEGL 243
QY 190 MKLITFTSVETVAKRVEKKAIVFGQEVNLSNDNEN-RKFDPSVASSLYKNWVPLT 248
DB 244 MQLITYETVERAIEHVEAKRIFGQEVNIGAKDNGSPVYKIRPSLVAELYSYRIMPLT 303
QY 249 KEVOVEYLLRLD 261
DB 304 KEVOVEYLLRLD 316
RESULT 12
ABU08095
ID ABU08095 standard; protein; 207 AA.
XX AC ABU08095;
XX AC ABU08095;

```

DT XX 22-MAY-2003 (first entry)  
DE XX Wheat chorismate mutase protein.  
XX Wheat; plant; enzyme; chorismate mutase; aromatic amino acid; diet;  
KW biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;  
KW phenylalanine; prephenate dehydrogenase; herbicide; food crop.  
XX OS Triticum aestivum.  
XX PN US2002184658-A1.  
XX PD 05-DEC-2002.  
XX PF 03-DEC-1999; 99US-00454279.  
XX PR 04-DEC-1998; 98US-0110845P.  
XX PA (FALC/) FALCO S C.  
XX PA (FAMO/) FAMODU O O.  
XX PA (LEEJ/) LEE J.  
XX PI Falco SC, Famodu OO, Lee J;  
XX DR WPI; 2003-328651/31.  
XX DR N-PSDB; ABX93059.  
XX PT New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
PT nucleic acids, useful for facilitating design and/or identifying  
PT inhibitors of those enzymes that may be used as herbicides and for  
PT producing antibodies.  
XX PS Claim 11; Fig 1; 32pp; English.  
XX CC The invention discloses isolated polynucleotides encoding chorismate  
CC mutase polypeptides. Aromatic amino acids must be included in the diets  
CC of animals. In the aromatic amino acid biosynthetic pathway chorismate is  
CC converted to anthranilate during tryptophan biosynthesis and is converted  
CC to prephenate, the branch point for tyrosine and phenylalanine  
CC biosynthesis. Chorismate mutase catalyses the conversion of chorismate to  
CC prephenate. Also disclosed are methods for selection of an isolated  
CC polynucleotide that affects the level of expression of a tyrosine  
CC biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
CC fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
CC evaluating compounds for their ability to inhibit the activity of a  
CC tyrosine biosynthetic enzyme. The polypeptides can be used produce  
CC antibodies. Chorismate mutase and prephenate dehydrogenase are good  
CC targets for herbicides that will not affect animals, and overexpression  
CC of these enzymes may be used to increase the content of aromatic amino  
CC acid in food crops. The polypeptides may also be used to design and/or  
CC identify inhibitors of those enzymes that may be used as herbicides. The  
CC nucleic acids may be used to create transgenic plants, as probes for the  
CC genetic and physical mapping of the genes and as markers for traits  
CC linked to those genes. The sequence presented is a wheat chorismate  
CC mutase protein  
XX Sequence 207 AA;  
SQ Query Match 45.0%; Score 602; DB 6; Length 207;  
Best Local Similarity 57.8%; Pred. No. 3.2e-55;  
Matches 118; Conservative 27; Mismatches 53; Indels 6; Gaps 2;  
QY 64 EFVVKNTAEIQAAGRYKXNPENAEFFPENLPPSVPSYKQFLHPCGAASINIKSTWKM 123  
DB 4 EFFVRAEVLHAKAGHYKQPEDVFFFSQDLPPVPPTKGRPKVLPFASLVCYNDALWKM 63  
QY 124 YFKLLPILATSGDGNVYAQTANDLSLQSIIRHYGKFAVSKFRDAPQDYELPIRA 183  
DB 64 YFNELLPLFTAGDGDGNVYAETVALDPAQLQALSRIHCGKYAEVKFKDAPQDYSPIRA 123  
QY 184 KDKGLMKLLTFTSVETVRKRVKKKAVVFGQENL----NSDDND--NENRKFDSVAS 237  
DB 124 KDTNALMNLITFTAVEKKVKRVEKKARIFGQNTLTEDSVGKQDGDACDCHCKVDKVL 183

QY 238 SLYKNWVPLTKVEVQVEYLLRLD 261  
DB 184 KLYDMWVPLTKDVEVEYLLRLD 207

RESULT 13  
AAG06003  
ID AAG06003 standard; protein; 319 AA.  
XX AC AAG06003;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 2617.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P;  
PR 05-MAR-1999; 99US-0123180P;  
PR 03-MAR-1999; 99US-0123548P;  
PR 23-MAR-1999; 99US-0125788P;  
PR 25-MAR-1999; 99US-0126264P;  
PR 29-MAR-1999; 99US-0126785P;  
PR 01-APR-1999; 99US-0127462P;  
PR 06-APR-1999; 99US-0128234P;  
PR 08-APR-1999; 99US-0128714P;  
PR 18-APR-1999; 99US-0129845P;  
PR 19-APR-1999; 99US-0130077P;  
PR 21-APR-1999; 99US-0130449P;  
PR 23-APR-1999; 99US-0130510P;  
PR 23-APR-1999; 99US-0130891P;  
PR 28-APR-1999; 99US-0131449P;  
PR 30-APR-1999; 99US-0132048P;  
PR 30-APR-1999; 99US-0132407P;  
PR 04-MAY-1999; 99US-0132484P;  
PR 05-MAY-1999; 99US-0132485P;  
PR 06-MAY-1999; 99US-0132486P;  
PR 07-MAY-1999; 99US-0132487P;  
PR 11-MAY-1999; 99US-0132863P;  
PR 14-MAY-1999; 99US-0134218P;  
PR 14-MAY-1999; 99US-0134219P;  
PR 14-MAY-1999; 99US-0134221P;  
PR 14-MAY-1999; 99US-0134370P;  
PR 18-MAY-1999; 99US-0134768P;  
PR 19-MAY-1999; 99US-0134941P;  
PR 20-MAY-1999; 99US-0135124P;  
PR 21-MAY-1999; 99US-0135353P;  
PR 24-MAY-1999; 99US-0135629P;  
PR 25-MAY-1999; 99US-0136021P;  
PR 27-MAY-1999; 99US-0136392P;  
PR 28-MAY-1999; 99US-0136782P;  
PR 01-JUN-1999; 99US-0137222P;  
PR 03-JUN-1999; 99US-0137528P;  
PR 04-JUN-1999; 99US-0137502P;  
PR 07-JUN-1999; 99US-0137724P;  
PR 08-JUN-1999; 99US-0138094P;  
PR 10-JUN-1999; 99US-0138540P;  
PR 10-JUN-1999; 99US-0138847P;  
PR 14-JUN-1999; 99US-0139119P;  
PR 16-JUN-1999; 99US-0139452P;  
PR 16-JUN-1999; 99US-0139453P;

PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 30-AUG-1999;	99US-0151080P.
PR 21-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151030P.
PR 22-JUN-1999;	99US-0139899P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140353P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140354P.	PR 07-SEP-1999;	99US-0152383P.
PR 24-JUN-1999;	99US-0140695P.	PR 10-SEP-1999;	99US-0153070P.
PR 28-JUN-1999;	99US-0140823P.	PR 13-SEP-1999;	99US-0153758P.
PR 29-JUN-1999;	99US-0140991P.	PR 15-SEP-1999;	99US-0154018P.
PR 30-JUN-1999;	99US-0141287P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0141842P.	PR 20-SEP-1999;	99US-0154779P.
PR 01-JUL-1999;	99US-0142154P.	PR 22-SEP-1999;	99US-0155113P.
PR 02-JUL-1999;	99US-0142055P.	PR 23-SEP-1999;	99US-0155486P.
PR 06-JUL-1999;	99US-0142390P.	PR 24-SEP-1999;	99US-0155659P.
PR 08-JUL-1999;	99US-0142803P.	PR 28-SEP-1999;	99US-0156458P.
PR 09-JUL-1999;	99US-0142920P.	PR 29-SEP-1999;	99US-0156596P.
PR 12-JUL-1999;	99US-0142977P.	PR 04-OCT-1999;	99US-0157117P.
PR 13-JUL-1999;	99US-0143542P.	PR 05-OCT-1999;	99US-0157533P.
PR 14-JUL-1999;	99US-0143624P.	PR 06-OCT-1999;	99US-0157865P.
PR 15-JUL-1999;	99US-0144005P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144085P.	PR 08-OCT-1999;	99US-0158232P.
PR 17-JUL-1999;	99US-0144086P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144325P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144333P.	PR 14-OCT-1999;	99US-0159339P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144352P.	PR 18-OCT-1999;	99US-0159584P.
PR 20-JUL-1999;	99US-0144632P.	PR 18-OCT-1999;	99US-0160741P.
PR 20-JUL-1999;	99US-0144884P.	PR 21-OCT-1999;	99US-0160767P.
PR 21-JUL-1999;	99US-0144814P.	PR 21-OCT-1999;	99US-0160768P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160770P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160814P.
PR 21-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160815P.
PR 22-JUL-1999;	99US-0145087P.	PR 22-OCT-1999;	99US-0160980P.
PR 22-JUL-1999;	99US-0145089P.	PR 22-OCT-1999;	99US-0160981P.
PR 22-JUL-1999;	99US-0145132P.	PR 22-OCT-1999;	99US-0160989P.
PR 23-JUL-1999;	99US-0145145P.	PR 25-OCT-1999;	99US-0161404P.
PR 23-JUL-1999;	99US-0145218P.	PR 25-OCT-1999;	99US-0161405P.
PR 23-JUL-1999;	99US-0145224P.	PR 25-OCT-1999;	99US-0161406P.
PR 26-JUL-1999;	99US-0145276P.	PR 26-OCT-1999;	99US-0161359P.
PR 27-JUL-1999;	99US-0145913P.	PR 26-OCT-1999;	99US-0161360P.
PR 27-JUL-1999;	99US-0145918P.	PR 26-OCT-1999;	99US-0161361P.
PR 27-JUL-1999;	99US-0145919P.	PR 28-OCT-1999;	99US-0161920P.
PR 28-JUL-1999;	99US-0145931P.	PR 28-OCT-1999;	99US-0161922P.
PR 02-AUG-1999;	99US-0146386P.	PR 28-OCT-1999;	99US-0161992P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161993P.
PR 03-AUG-1999;	99US-0146389P.	PR 29-OCT-1999;	99US-0162142P.
PR 04-AUG-1999;	99US-0147038P.		
PR 04-AUG-1999;	99US-0147204P.		
PR 04-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		

Query Match 44.8%; Score 599.5; DB 3; Length 319;

Best Local Similarity 48.7%; Pred. No. 1.2e-54;

Matches 113; Conservative 45; Mismatches 71; Indels 3; Gaps 1;

QY 10 DSGNVTLASVREDLVROEDTIIYGLIERAKPFSNHYDEKIAQIOFCGSLVFEVKN 69

Db 75 DESESLTLEGIRNSLIROEDSIIFGLIERAKYCYNADYDPTAFDMDFGNSLVEYMWKG 134

QY 70 TEAIQAKAGRYKNPBNAPFFENLPPSIVPSYSEKQFLHPGAASININKSIWYFKELL 129

Db 135 TELHAKVGRFKSPDEHPFDDLPBPMLPPLQYPKVLHFAADSININKKIWNMYFRDIV 194  
QY 130 PLLATSGDDGNVAQTAANDLSLQISRIHGVKFTVAEYKFRDAPQDYPLIRAKDKGL 189  
Db 195 PRUVKGGDGNVGTAVCDALCLQCLSKRIHYGFVAEAKFQASPEAYESAIAKQDKDAL 254  
QY 190 MKLLTFTSVEETVRKEVEKAVVFGQEVNLSDDNDNENRKPDPSPVASSLYK 241  
Db 255 MDMLTPTVEDAIKKRVEKTRTYGQEVKVGMEKEEBEEBEGN---ESHVYK 303

RESULT 14  
AAG06002  
ID AAG06002 standard; protein; 325 AA.  
XX AC AAG06002;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 2616.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136382P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 08-JUL-1999; 99US-0142350P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 12-JUL-1999; 99US-0142920P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148585P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 21-OCT-1999; 99US-0159594P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160993P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161932P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 44.8%; Score 599.5; DB 3; Length 325;  
Best Local Similarity 48.7%; Pred. No. 1.2e-54;  
Matches 113; Conservative 45; Mismatches 71; Indels 3; Gaps 1;

QY 10 DSGNVYTLASVREDLVQEDTIYGLERAKFFSNHSHYDEKYAQIQGFCGSLVEFVKN 69

Db 81 DESESLTEGIRNSLRQDSIIIFGLERAKYCYNADTYDTAFDMDFGNSLVEYVWG 140  
QY 70 TEAIQAKAGRYKPNENAFPPENLPSPISVPSYFKQFLHPGAASININKSINKMYEKELL 129  
Db 141 TEKHAQVGRFKSPDEHPFFDLPPEMLPPQYQPKVLHFAADSNINKKINWNYFRDLV 200  
QY 130 PLLATSGDDGNYAQTAAANDLSLIQSISSRIHYGKFVAEVKFRDPAQDYELIRAKDK 189  
Db 201 PRLVKGDDGNYGSTAVCDALICQLCKRIHYGKFVAEAKFQASPEAVESAIAQDKDAL 260  
QY 190 MKLLTTSVEETVRKEVEKAVVFGGVNLNSDDNDNENKRPDPSPVASSLYK 241  
Db 261 MDMLTFTVEDAIKRVEMKTRTYGQEVKVGMEKEEEEEEGN---ESHVYK 309

## RESULT 15

AAG06001  
ID AAG06001 standard; protein; 347 AA.

XX AAG06001;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 2615.

XX Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 09-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 16-APR-1999; 99US-0128714P.

XX 19-APR-1999; 99US-0129845P.

XX 21-APR-1999; 99US-0130077P.

XX 23-APR-1999; 99US-0130510P.

XX 28-APR-1999; 99US-0130891P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 30-APR-1999; 99US-0132407P.

XX 04-MAY-1999; 99US-0132484P.

XX 05-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 07-MAY-1999; 99US-0132487P.

XX 11-MAY-1999; 99US-0132863P.

XX 11-MAY-1999; 99US-0134256P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 14-MAY-1999; 99US-0134370P.

XX 18-MAY-1999; 99US-0134768P.

XX 19-MAY-1999; 99US-0134941P.

XX 20-MAY-1999; 99US-0135124P.

XX 21-MAY-1999; 99US-0135353P.

XX 24-MAY-1999; 99US-0135629P.

XX 25-MAY-1999; 99US-0136021P.

XX 27-MAY-1999; 99US-0136392P.

XX 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142058P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144008P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145132P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157573P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0158293P.  
PR 13-OCT-1999; 99US-0158294P.  
PR 13-OCT-1999; 99US-0158295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 25-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 44.8%; Score 599.5; DB 3; Length 347;  
Best Local Similarity 48.7%; Pred. No. 1.3e-54;  
Matches 113; Conservative 45; Mismatches 71; Indels 3; Gaps 1;  
QY 10 DSGNVTIASVREDIVROEDTIIYGLIRAKFPNSHTYDEKYAIOGFGSLVEFYKN 69  
Db 103 DESESLITLEGIRNSLIRQEDSLIFGLERAKYCYNADTYDPTAPDMDFNGSLVEYKVG 162  
QY 70 TEAIOAKAGRYKNPEENAFPPENLPSPISVSFKQFLHPGASININKSIWKMYFKEL 129  
Db 163 TEKHAAGRGKSPDEHPFPDPDLPEPMLPQYKVLHPAADSININKIWMYFRDLV 222  
QY 130 PILATSGDDGNYAQTANDLSLISISRIHYGKFAVAFKPRDAQDYEPILIRAKDEGL 189  
Db 223 PLYVKKGGDGNYSTAVCDALICLQCLSKRIHYGKFAEAKFQASPEAYESAIKQDKDAL 282  
QY 190 KMLTFETSVETVRKREKKAIVFGQEVNLNSDNDNENNRKFPDSVASLYK 241  
Db 283 MDVLTFFIVEDAIXKRVEMKRTYQGEYKVGMEKEKEEBEEN--ESHVYK 331

Search completed: September 25, 2004, 02:52:38  
Job time : 77 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004. Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2004, 02:51:21; Search time 23 seconds  
(without alignments)  
585.843 Million cell updates/sec

Title: US-10-624-061-16  
Perfect score: 1339  
Sequence: 1 MAKAAEQSPDSGNVYTLASV.....NWVPLTKREVQVYLLRLD 261

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues  
Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pep:\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pep:\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pep:\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pep:\*  
5: /cgn2\_6/prodata/2/iaa/PCUTS COMB.pep:\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	1339	100.0	261	4	US-09-454-279-16
2	818	61.1	265	4	US-09-454-279-22
3	696	52.0	334	4	US-09-454-279-21
4	681.5	50.9	257	4	US-09-454-279-14
5	673	50.3	132	4	US-09-454-279-6
6	670.5	50.1	331	4	US-09-454-279-12
7	602	45.0	207	4	US-09-454-279-18
8	489	36.5	154	4	US-09-454-279-8
9	147	11.0	93	4	US-09-454-279-2
10	131	9.8	62	4	US-09-454-279-4
11	96	7.2	464	4	US-09-134-001C-4701
12	96	7.2	753	4	US-09-543-681A-5022
13	92	6.9	946	3	US-09-074-579-3
14	92	6.9	946	3	US-09-388-774-3
15	90.5	6.8	554	3	US-09-319-989-6
16	89.5	6.7	590	4	US-09-107-532A-4682
17	89	6.5	808	4	US-09-543-681A-4995
18	86.5	6.5	930	4	US-09-134-001C-5314
19	86.5	6.5	930	4	US-09-386-962C-10
20	83.5	6.2	1138	1	US-09-973-320-2
21	83	6.2	612	1	US-08-344-695-2
22	83	6.2	769	4	US-09-540-236-3110
23	83	6.2	902	4	US-09-107-532A-3837
24	83	6.2	1070	2	US-08-633-770A-2
25	83	6.2	1070	4	US-08-280-197-6
26	81	6.0	851	4	US-09-071-035-326
27	81	6.0	851	4	US-09-071-035-330

28	81	6.0	851	4	US-09-071-035-334	Sequence 334, Appl
29	81	6.0	946	4	US-09-546-153-1	Sequence 1, Appli
30	81	6.0	962	4	US-09-134-000C-4257	Sequence 4257, Ap
31	80.5	6.0	1180	3	US-09-224-024-28	Sequence 28, Appl
32	80.5	6.0	1180	5	PCT-US94-07902-28	Sequence 28, Appl
33	80.5	6.0	1440	3	US-09-357-251-37	Sequence 37, Appl
34	80.5	6.0	1512	4	US-09-443-184-48	Sequence 48, Appl
35	80	6.0	1110	3	US-08-118-441-29	Sequence 29, Appl
36	80	6.0	1110	3	US-08-338-579A-29	Sequence 29, Appl
37	80	6.0	1110	5	PCT-US94-09851-29	Sequence 29, Appl
38	79.5	5.9	438	4	US-09-107-532A-3764	Sequence 3764, Ap
39	79.5	5.9	535	4	US-08-969-046-2	Sequence 2, Appli
40	79.5	5.9	998	2	US-08-415-788-7	Sequence 7, Appli
41	79.5	5.9	1029	2	US-08-415-788-3	Sequence 3, Appli
42	79.5	5.9	1792	4	US-09-561-818A-12	Sequence 12, Appl
43	79.5	5.9	1816	4	US-09-561-818A-10	Sequence 10, Appl
44	79	5.9	997	1	US-08-232-540-1	Sequence 1, Appli
45	79	5.9	997	1	US-08-428-949A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-454-279-16  
; Sequence 16, Application US/09454279  
; Patent No. 6627798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BBI299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454, 279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 16  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-454-279-16

Query Match	100.0%;	Score 1339;	DB 4;	Length 261;
Best Local Similarity	100.0%;	Pred. No. 2e-137;		
Matches 261;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAKAAEQSPDSGNVYTLASVREDLV	QEDTIIYGLIERAKFPNSHTYDEKYAQIQGFCG	60
Db	1	MAKAAEQSPDSGNVYTLASVREDLV	QEDTIIYGLIERAKFPNSHTYDEKYAQIQGFCG	60
Qy	61	SLVEFVVKNTAIQAKAGRYKNPEENAFPPENLPSIVPSYSFQKFLHPGGAASINIKSI	120	
Db	61	SLVEFVVKNTAIQAKAGRYKNPEENAFPPENLPSIVPSYSFQKFLHPGGAASINIKSI	120	
Qy	121	WKVYFKELIPLLATSGDDGNYAQTAAANDLSLQSTISRIHYGKFVAEVKFRDAPQDYEP	180	
Db	121	WKVYFKELIPLLATSGDDGNYAQTAAANDLSLQSTISRIHYGKFVAEVKFRDAPQDYEP	180	
Qy	181	IRAKDKGLMKLLTFTSVETVRKVEKKAQVVFQGVNLNSDDNDNENRKDFPDSVASSLY	240	
Db	181	IRAKDKGLMKLLTFTSVETVRKVEKKAQVVFQGVNLNSDDNDNENRKDFPDSVASSLY	240	
Qy	241	KNWVPLTKREVQVEYLLRLD	261	
Db	241	KNWVPLTKREVQVEYLLRLD	261	

RESULT 2  
US-09-454-279-22  
; Sequence 22, Application US/09454279

Patent No. 6627798  
 GENERAL INFORMATION:  
 APPLICANT: Falco, S. Carl  
 APPLICANT: Famodu, Omolayo O.  
 APPLICANT: Lee, Jian-Ming  
 TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 FILE REFERENCE: BB1299 US NA  
 CURRENT APPLICATION NUMBER: US/09/454,279  
 CURRENT FILING DATE: 1999-12-03  
 EARLIER APPLICATION NUMBER: 60/110,845  
 EARLIER FILING DATE: 1998-12-04  
 NUMBER OF SEQ ID NOS: 22  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 22  
 TYPE: PRT  
 ORGANISM: Arabidopsis thaliana  
 US-09-454-279-22

Query Match 61.1%; Score 818; DB 4; Length 265;  
 Best Local Similarity 61.1%; Pred. No. 1e-80;  
 Matches 162; Conservative 31; Mismatches 68; Indels 4; Gaps 2;  
 QY 1 MAKAAEQSDSG--NVYTLASVREDLVROEDTIIYGLIERAKFPSNSHTYDEKVAQIQGF 58  
 Db 1 MARVFESDGSNCNLSLDLRESLIRQEDTIVFSLIERAKFPLNSPAPFESRCLDSGS 60  
 QY 59 CGSLVEFVVNTEAIQAKAGRYKNEENAFPPENLPPSIVPSYFKQFLHPGAASININK 118  
 Db 61 FSLTEFFVRETEIIQAKGRVEYEENPFLENIPHSVFTKYPSALHPKALSVINK 120  
 QY 119 STWKYFELLPLATSGDGNAGTAANDLSLQSIERRHYGKFAEVRPDAPODYE 178  
 Db 121 QIWDIYFELLPLFVKPGDGNYPSTAASDLACLQALSRIHYGKFAEVRPDAPODYE 180  
 QY 179 PLIRAKDEGLMKLITFTSVETVRKVEKAVVFGQVWLNSSDDNNEKFE--DPSVA 236  
 Db 181 PAIRAQDREALMKLITFEKVEWVKRQKAEIFGQVKNKSGYKDESKKYYKVDPLLA 240  
 QY 237 SLYKNWVPLTKVEQVYLLRLD 261  
 Db 241 SRIYGEWLIPLTKVEVYLLRLD 265

RESULT 3  
 US-09-454-279-21  
 Query Match 52.0%; Score 696; DB 4; Length 334;  
 Best Local Similarity 50.4%; Pred. No. 2.8e-67;  
 Matches 131; Conservative 46; Mismatches 75; Indels 8; Gaps 1;  
 QY 10 DSGNVYTLASVREDLVROEDTIIYGLIERAKFPSNSHTYDEKVAQIQGFGSLVEFVKN 69  
 Db 75 DESSELTLEGIRNSLIRQEDSIIFGLIERAKYCYNADTYDPTAFDWDGNGSLVEFVKN 134  
 GENERAL INFORMATION:  
 APPLICANT: Falco, S. Carl  
 APPLICANT: Famodu, Omolayo O.  
 APPLICANT: Lee, Jian-Ming  
 TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 FILE REFERENCE: BB1299 US NA  
 CURRENT APPLICATION NUMBER: US/09/454,279  
 CURRENT FILING DATE: 1999-12-03  
 EARLIER APPLICATION NUMBER: 60/110,845  
 EARLIER FILING DATE: 1998-12-04  
 NUMBER OF SEQ ID NOS: 22  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 21  
 TYPE: PRT  
 ORGANISM: Arabidopsis thaliana  
 US-09-454-279-21

QY 70 TEAIQAKAGRYKNEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWMPFKELL 129  
 Db 135 TEKLHAKVGRKSPDBHPPFPDLPBPMLPPOYKVLHFAADSIININKIWNMYFRDIV 194  
 QY 130 PLLATSGDGNAGTAANDLSLQSIERRHYGKFAEVRPDAPODYEPLIRAKDKGL 189  
 Db 195 PRLVKGGDGNAGTAVCAICLQSLKRIHYGKFAEVRPDAPODYEPLIRAKDKGL 254  
 QY 190 MKLLTFTSVETVRKVEKAVVFGQVWLNSSDDNNEKFE--KFDPSVASSLYK 241  
 Db 255 MDMLTFTPTVEDAIKRVEMKTRTYGQEVKVGMEKEEKEEENSHVYKISILVGLDLYG 314  
 QY 242 NWVPLTKVEQVYLLRLD 261  
 Db 315 DWIMPLTKVEQVYLLRLD 334

RESULT 4  
 US-09-454-279-14  
 Query Match 50.9%; Score 681.5; DB 4; Length 257;  
 Best Local Similarity 52.1%; Pred. No. 7e-66;  
 Matches 134; Conservative 44; Mismatches 78; Indels 1; Gaps 1;  
 QY 6 EOPSGNVYTLASVREDLVROEDTIIYGLIERAKFPSNSHTYDEKVAQIQGFGSLVEF 65  
 Db 1 KORIDQSEILTDNIRTSILVRQEDSIIFSLIERAQFCYNADIYDKNAPHVDFGDSLVEF 60  
 QY 66 VVKNTAEIQAAGRYKNEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWMPYF 125  
 Db 61 WURETEKLUHQVGRYKSPDEHPFPELPLPQYKVLHPIADSIININKSIWMPYF 120  
 QY 126 KELLPLATSGDGNAGTAANDLSLQSIERRHYGKFAEVRPDAPODYEPLIRAKD 185  
 Db 121 DELLPLRVKESDGNYSALCDTICLQALSRIHYGKFAEVRPDAPODYEPLIRAKD 180  
 QY 186 KEGLMKLLTFTSVETVRKVEKAVVFGQVWLNSSDDNNEKFE--KFDPSVASSLYK 244  
 Db 181 CDLMHLLTYETVERAIEHRVEKAKIFGQVLDIAGDNGAPPYKIRPSLVABEYSYRI 240  
 QY 245 IPLTKVEQVYLLRLD 261  
 Db 241 MELTKVEQVYLLRLD 257  
 GENERAL INFORMATION:  
 APPLICANT: Falco, S. Carl  
 APPLICANT: Famodu, Omolayo O.  
 APPLICANT: Lee, Jian-Ming  
 TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 FILE REFERENCE: BB1299 US NA  
 CURRENT APPLICATION NUMBER: US/09/454,279  
 CURRENT FILING DATE: 1999-12-03  
 EARLIER APPLICATION NUMBER: 60/110,845  
 EARLIER FILING DATE: 1998-12-04  
 NUMBER OF SEQ ID NOS: 22  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 14  
 LENGTH: 257  
 TYPE: PRT  
 ORGANISM: Oryza sativa  
 US-09-454-279-14

Query Match 50.9%; Score 681.5; DB 4; Length 257;  
 Best Local Similarity 52.1%; Pred. No. 7e-66;  
 Matches 134; Conservative 44; Mismatches 78; Indels 1; Gaps 1;  
 QY 6 EOPSGNVYTLASVREDLVROEDTIIYGLIERAKFPSNSHTYDEKVAQIQGFGSLVEF 65  
 Db 1 KORIDQSEILTDNIRTSILVRQEDSIIFSLIERAQFCYNADIYDKNAPHVDFGDSLVEF 60  
 QY 66 VVKNTAEIQAAGRYKNEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWMPYF 125  
 Db 61 WURETEKLUHQVGRYKSPDEHPFPELPLPQYKVLHPIADSIININKSIWMPYF 120  
 QY 126 KELLPLATSGDGNAGTAANDLSLQSIERRHYGKFAEVRPDAPODYEPLIRAKD 185  
 Db 121 DELLPLRVKESDGNYSALCDTICLQALSRIHYGKFAEVRPDAPODYEPLIRAKD 180  
 QY 186 KEGLMKLLTFTSVETVRKVEKAVVFGQVWLNSSDDNNEKFE--KFDPSVASSLYK 244  
 Db 181 CDLMHLLTYETVERAIEHRVEKAKIFGQVLDIAGDNGAPPYKIRPSLVABEYSYRI 240  
 QY 245 IPLTKVEQVYLLRLD 261  
 Db 241 MELTKVEQVYLLRLD 257

RESULT 5  
 US-09-454-279-6  
 Query Match 52.0%; Score 696; DB 4; Length 334;  
 Best Local Similarity 50.4%; Pred. No. 2.8e-67;  
 Matches 131; Conservative 46; Mismatches 75; Indels 8; Gaps 1;  
 QY 10 DSGNVYTLASVREDLVROEDTIIYGLIERAKFPSNSHTYDEKVAQIQGFGSLVEFVKN 69  
 Db 75 DESSELTLEGIRNSLIRQEDSIIFGLIERAKYCYNADTYDPTAFDWDGNGSLVEFVKN 134  
 GENERAL INFORMATION:  
 APPLICANT: Falco, S. Carl  
 APPLICANT: Famodu, Omolayo O.  
 APPLICANT: Lee, Jian-Ming

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; FILE REFERENCE: BB1299 US NA

; CURRENT APPLICATION NUMBER: US/09/454,279

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110,845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 6

; LENGTH: 132

; TYPE: PRT

; ORGANISM: Glycine max

; US-09-454-279-6

Query Match 50.3%; Score 673; DB 4; Length 132;

Best Local Similarity 98.5%; Pred. No. 2.1e-65;

Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 DSGNVYTLASVREDLVQRQEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCGSLVEFVVK 69

DB 1 DSGNVYTLASVREDLVQRQEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCGSLVEFVVK 60

QY 70 TEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWKYFKELL 129

DB 61 TEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWKYFKELL 120

QY 130 PLLATSGDGGNY 141

DB 121 PLLATSGDGGNY 132

#### RESULT 6

US-09-454-279-12

; Sequence 12, Application US/09454279

; Patent No. 6627798

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Famodu, Omolayo O.

; APPLICANT: Lee, Jian-Ming

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; FILE REFERENCE: BB1299 US NA

; CURRENT APPLICATION NUMBER: US/09/454,279

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110,845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 12

; LENGTH: 331

; TYPE: PRT

; ORGANISM: Zea mays

; US-09-454-279-12

Query Match

Best Local Similarity 50.1%; Score 670.5; DB 4; Length 331;

Matches 130; Conservative 48; Mismatches 74; Indels 1; Gaps 1;

QY 10 DSGNVYTLASVREDLVQRQEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCGSLVEFVVK 69

DB 79 DRSDILFDSIRQVLINLEDSIIIFGLIERAQFCYNADTYDSNAFMDGFGGSLVEIVRE 138

QY 70 TEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWKYFKELL 129

DB 139 TEKLHAQVGRYKSPDEHPFESKDLPEPLPPMOYPRVLHPPIADSININKSIWKYFDELL 198

QY 130 PLLATSGDGGNYAQTAANDLSLQSRRIHYGKFAVEKFRDAPQDYELPIRAKDEGL 189

DB 199 FRLVKESDGNAGSSALCDTCTCLQALSRIHYGKFAVEAKFQSPAYTPAIIAQDRDQL 258

QY 190 MKLLTFTSVETVRKVEKAVVFGQVNLNSDDNEN-RKDPDSVASSLYKMWIPLT 248

DB 259 MNLVTETVERAEHRVEAKAKIFGQEVNICKDNGSPVYKIRPSLVAEIYSRIMPLT 318

QY 249 KEVQVEYLLRLD 261

DB 319 KEVEYALLRLD 331

#### RESULT 7

US-09-454-279-18

; Sequence 18, Application US/09454279

; Patent No. 6627798

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Famodu, Omolayo O.

; APPLICANT: Lee, Jian-Ming

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; FILE REFERENCE: BB1299 US NA

; CURRENT APPLICATION NUMBER: US/09/454,279

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110,845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 18

; LENGTH: 207

; TYPE: PRT

; ORGANISM: Triticum aestivum

; US-09-454-279-18

Query Match

Best Local Similarity 45.0%; Score 602; DB 4; Length 207;

Matches 118; Conservative 27; Mismatches 53; Indels 6; Gaps 2;

QY 64 EFVKNTEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWK 123

DB 4 EFFVREAEVLHAKAGHYQKPEDVPFSSODLPPVPFTKGRPKVLHPFASLVCVNDALWK 63

QY 124 YFKELLPLATSGDGGNYAQTAANDLSLQSRRIHYGKFAVEKFRDAPQDYELPIRA 183

DB 64 YFNEILLPLTADGDDGNYAETVALDFACLOALSRIHYGKFAVEKFRDAPQDYSPPIRA 123

QY 184 KDKGKMLLTFTSVETVRKVEKAVVFGQVNLNSDDND--NENRKEPDSVAS 237

DB 124 KDTNALMNLFTAVEEKKVEKKEKARIFQGNVTLEDSVGKQDQDADCSHCKVDPKVLS 183

QY 238 SLYKNWVPLTKVEQVYLLRLD 261

DB 184 KLYDMWVPLTKDVEYVLLRLD 207

#### RESULT 8

US-09-454-279-8

; Sequence 8, Application US/09454279

; Patent No. 6627798

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Famodu, Omolayo O.

; APPLICANT: Lee, Jian-Ming

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; FILE REFERENCE: BB1299 US NA

; CURRENT APPLICATION NUMBER: US/09/454,279

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110,845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 8

; LENGTH: 154

; TYPE: PRT

; ORGANISM: Triticum aestivum

; US-09-454-279-8

Query Match

Best Local Similarity 36.5%; Score 489; DB 4; Length 154;

Matches 92; Conservative 22; Mismatches 40; Indels 0; Gaps 0;

QY	64	EFVVKNTETAIQAKAGRYKNPDENAFFPENLPSPVPSFKQFLHPCAASININKSIWKM	123
Db	1	EFFVREAEVLHAKAGHYKQKPEDVFFSQDLPPVPPTKGRPKVLHPFASIVCVNDAIWKM	60
QY	124	YFKEILLPLLATSGDDGNVAQTAANDLSLQISRIHYGKFAEVEVKFRAPADQDYELPIRA	183
Db	61	YFNEILLPFTADGDGNVTAETVALDFCQLQALSRHICGKYFAEVEVKFADQDYSEPIRA	120
QY	184	KDKGIMKLLTFTSYVEETVRKREKKAUVFGQEV	217
Db	121	KDTNALMLLTFTAVEEKVKRVEKKKARIFGQNV	154

```

RESULT 9
US-09-454-279-2
; Sequence 2, Application US/09454279
; Patent No. 6627798
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/09/454, 279
; CURRENT FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
US-09-454-279-2

```

RESULT 10  
US-09-454-279-4  
Sequence 4, Application US/09454279  
Patent No. 6627798  
GENERAL INFORMATION:  
APPLICANT: Falco, S. Carl  
APPLICANT: Famodu, Omlayo O.  
APPLICANT: Lee, Jian-Ming  
TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
FILE REFERENCE: BB1299 US NA  
CURRENT APPLICATION NUMBER: US/09/454,279  
CURRENT FILING DATE: 1999-12-03  
EARLIER APPLICATION NUMBER: 60/110,845  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 4  
LENGTH: 62  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (49)  
FEATURE:

```
; NAME/KEY : UNSURE  
; LOCATION : (58)  
US-09-454-279-4
```

	Query Match	9.8%;	Score 131;	DB 4;	Length 62;
	Best Local Similarity	41.9%;	Pred. No.	6,1e-07;	
	Matches	26;	Conservative	14;	Mismatches 22;
					Indels
QY	6	EQSPDGNVYTTLASVRDLVRQEDTIIVGLIERAKFPNSNHYDEKIAQ	:	:	:
		: :	:	:	:
Dd	1	KQRICQEITLTDNRISLVRCQDSIIIFSLLERAQFATNADIYDKNAFX	:	:	:
		: :	:	:	:
QY	66	VV	67		
	:	:	:		
Dd	61	MV	62		

```

RESULT 11
US-09-134-001C-4701
; Sequence 4701, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: LYNN Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4701
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4701

```

RESULT 12  
US-09-543-681A-5022  
; Sequence 5022, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEOTIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001

# CONSTITUTIONAL LAW

NAME

REC

REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0505 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 946 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENE BANK  
CLONE: gi33985  
US-09-388-774-3

Query Match 6.9%; Score 92; DB 3; Length 946;  
Best Local Similarity 19.2%; Pred. No. 0.78;  
Matches 56; Conservative 43; Mismatches 89; Indels 104; Gaps 13;

QY 19 SVREDLVQRQDIYGLIERAKPPNSHTYDEKVAIQGFCGSLVEFVVKNTAEIAQAKAG 78  
Db 263 SRETAVDGLVLYD-VKR-----BEKAGELEVFNQYFVHF----- 298  
QY 79 RYKNPEENAFPPENLPSPISVPSYFKQFLHPGCAASININKSIWKMYFKELLPPLATSGDD 138  
Db 299 -----FAPDNLDP--IP-----KNILF-----VIDVSGSMGWGVMQKQTVEMAKTILDD 339  
QY 139 -----GNYAQTAANDLSLQSISSRIHYGK-FVAEVRFRDAPQDYEPFLIRA-- 183  
Db 340 LPAEDHFSVIDFNQIRWNLDFLOQ--KHLQIAKYIEKIQPSGGTINAEALLRAIF 397  
QY 184 -----KDEGLMKLLTFTSVETVRKVEKAVVFGO--EVLNLSDDNDNEN 228  
Db 398 ILNEANNLGLDPNSVSLIIVSDGDPVGBLKLSKIQKNVKNENIODNISLFSLGMGFDV 457  
QY 222 DDN-----DNEKRKDPSSVASSLYKNW-----VIPLTKVQVVEY 255  
Db 458 DYDFLKRSLNENH-----GIAQRIYQNTSSQLKFKFYNQVSTPLLRNVQFNY 505

RESULT 15  
US-09-319-989-6  
Sequence 6, Application US/09319989  
Patent No. 6130914  
GENERAL INFORMATION:  
APPLICANT: Grivell, Leslie A.  
APPLICANT: Teixeira De Mattos, Maarten J.  
APPLICANT: Biocm, Jolanda  
TITLE OF INVENTION: METHODS FOR MODULATING METABOLIC PATHWAYS OF  
TITLE OF INVENTION: MICRO-ORGANISMS AND MICRO-ORGANISMS OBTAINABLE BY SAID  
TITLE OF INVENTION: METHODS  
FILE REFERENCE: 24615-20123.00  
CURRENT APPLICATION NUMBER: US/09/319,989  
CURRENT FILING DATE: 1999-06-14  
EARLIER APPLICATION NUMBER: PCT/NL97/00688  
EARLIER FILING DATE: 1997-12-12  
EARLIER APPLICATION NUMBER: EPO 96203520  
EARLIER FILING DATE: 1996-12-12  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 6  
LENGTH: 554  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: HAP4  
US-09-319-989-6

Query Match 6.8%; Score 90.5; DB 3; Length 554;  
Best Local Similarity 22.2%; Pred. No. 0.49;  
Matches 51; Conservative 37; Mismatches 85; Indels 57; Gaps 11;

QY 37 ERAKPPNSHTYDEKVAIQGFCGSLVEFVVKNTAEIAQAKA-----GRYKNP 83  
Db 124 ERKIPRIHTIDEKLINDSNYL-AFLKFDLLENKFFHSSASSISPSYSPFSFSSYRNR 182  
QY 84 EENAFPPENLPSPISVPSYFKQFL---HPGAASININ-----KSIWKMYFKELLPPLATS 135  
Db 183 KKEFWDDSCDVTETIAHNSLLTKNHHIDSSNVHAPPTKKS--KLNDFDLLSLSTS 240  
QY 136 GDDGNYAQTAANDLSLQSISSRIHYGKFAEVRFRDAPQDYEP-----LIR----- 182  
Db 241 -----SSATVPQLTKDLANNLNFHFKIPHKASPPDSPADFPSPADSVSLIRNHSPLTNL 293  
QY 183 -AKDK-EGLMKLLTFTSVETVRKVEKAVVFGO--EVLNLSDDNDNEN 228  
Db 294 QVKDKIEDLNEIKFFNDFE-----KLEFFNKYAKVNTNNDYNEKND 334

Search completed: September 25, 2004, 02:55:14  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2004, 02:54:21 ; Search time 73 Seconds  
(without alignments)  
1149.683 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKAAEQSPDSGNVYTLASV.....NWVPLTKVQVEYLLRRLD 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCT05\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1339	100.0	261	US-09-454-279-16
2	1339	100.0	261	US-10-624-061-16
3	1339	100.0	297	US-10-425-114-48703
4	1315	98.2	261	US-10-424-599-235878
5	970	72.4	199	US-10-425-114-43606
6	818	61.1	265	US-09-454-279-22
7	818	61.1	265	US-10-624-061-22
8	696	52.0	334	US-09-454-279-21
9	685.5	51.2	313	US-10-624-061-21
10	681.5	50.9	257	US-10-437-963-196439
11	681.5	50.9	257	US-09-454-279-14
12	681.5	50.9	257	US-10-624-061-14
13	678.5	50.7	255	US-10-437-963-151855
14	673	50.3	132	US-09-454-279-6
15	673	50.3	132	US-10-624-061-6

16	671	50.1	264	12	US-10-425-114-66528	Sequence 66528, A
17	670.5	50.1	331	9	US-09-454-279-12	Sequence 12, Appl
18	668.5	50.1	331	15	US-10-624-061-12	Sequence 12, Appl
19	668.5	49.9	350	16	US-10-767-701-45917	Sequence 45917, A
20	661	48.6	320	12	US-10-425-114-68586	Sequence 68586, A
21	650	48.5	315	12	US-10-424-599-266862	Sequence 266862, A
22	636.5	47.5	313	12	US-10-425-114-65608	Sequence 65608, A
23	602	45.0	207	9	US-09-454-279-18	Sequence 18, Appl
24	602	45.0	207	15	US-10-624-061-18	Sequence 18, Appl
25	559	41.7	210	16	US-10-767-701-32934	Sequence 32934, A
26	543	40.6	316	16	US-10-437-963-186040	Sequence 186040, A
27	523	39.1	306	16	US-10-437-963-186047	Sequence 186047, A
28	515	38.5	489	16	US-10-437-963-103185	Sequence 103185, A
29	502.5	37.5	186	12	US-10-425-114-52677	Sequence 52677, A
30	489	36.5	154	9	US-09-454-279-8	Sequence 8, Appl
31	489	36.5	154	15	US-10-624-061-8	Sequence 8, Appl
32	443	33.1	109	12	US-10-424-599-224668	Sequence 224668, A
33	423	31.6	280	13	US-10-042-059A-2	Sequence 2, Appl
34	234.5	17.5	97	12	US-10-425-114-38176	Sequence 38176, A
35	188.5	14.1	83	12	US-10-425-114-52545	Sequence 52545, A
36	188.5	14.1	89	12	US-10-424-599-272410	Sequence 2, Appl
37	147	11.0	93	9	US-09-454-279-2	Sequence 4, Appl
38	147	11.0	93	15	US-10-624-061-2	Sequence 284420, A
39	140	10.5	118	12	US-10-424-599-276200	Sequence 11269, A
40	131	9.8	62	9	US-09-454-279-4	Sequence 58574, A
41	131	9.8	62	15	US-10-624-061-4	Sequence 1796, Ap
42	112	8.4	59	12	US-10-424-599-284420	
43	103	7.7	781	9	US-09-815-242-11269	
44	103	7.7	781	12	US-10-282-122A-58574	
45	95.5	7.1	1556	15	US-10-369-493-1796	

#### ALIGNMENTS

#### RESULT 1

US-09-454-279-16  
; Sequence 16, Application US/09454279  
; Publication No. US20020184658A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Pamodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 16  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-454-279-16

Query Match	100.0%;	Score 1339;	DB 9;	Length 261;
Best Local Similarity	100.0%;	Pred. No. 1.2e-125;		
Matches 261;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAKAAEQSPDSGNVYTLASVREDLVQRDIIIVGLIERAKFPNSHTYDKYAIQGFQCG	60	
Db	1	MAKAAEQSPDSGNVYTLASVREDLVQRDIIIVGLIERAKFPNSHTYDKYAIQGFQCG	60	
Qy	61	SLVEFVVKNTAEIAQAKAGRYKKNPEENAFFPENPPSVPSYFKQFLHPGAASININKSI	120	
Db	61	SLVEFVVKNTAEIAQAKAGRYKKNPEENAFFPENPPSVPSYFKQFLHPGAASININKSI	120	
Qy	121	WKMYFKELLPLLATSGDDGNVYTAQTAANDLSLQSIISRRIHYKXFAVKFRDPAQDYEP	180	
Db	121	WKMYFKELLPLLATSGDDGNVYTAQTAANDLSLQSIISRRIHYKXFAVKFRDPAQDYEP	180	

QY 181 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 240  
 DB 181 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 240  
 QY 241 KMWIPLTKVEQVEYLLRLD 261  
 DB 241 KMWIPLTKVEQVEYLLRLD 261

RESULT 2  
 US-10-624-061-16  
 ; Sequence 16, Application US/10624061  
 ; Publication No. US20040019929A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Falco, S. Carl  
 ; APPLICANT: Famodu, Omolayo O.  
 ; APPLICANT: Lee, Jian-Ming  
 ; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 ; FILE REFERENCE: BB1299 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/624,061  
 ; CURRENT FILING DATE: 2003-07-21  
 ; PRIOR APPLICATION NUMBER: US/09/454,279  
 ; PRIOR FILING DATE: 1999-12-03  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 16  
 ; LENGTH: 261  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 US-10-624-061-16

Query Match 100.0%; Score 1339; DB 15; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-125;  
 Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAKAAEQSPDSGNVYTLASVREDLVQEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCG 60  
 DB 1 MAKAAEQSPDSGNVYTLASVREDLVQEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCG 60  
 QY 61 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHFGAASININKSI 120  
 DB 61 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHFGAASININKSI 120  
 QY 121 WKMYFKELLPLLATSGDGNVYTAANDLSLQSSRIHYGKFAVVKFRDAPQDYEP 180  
 DB 121 WKMYFKELLPLLATSGDGNVYTAANDLSLQSSRIHYGKFAVVKFRDAPQDYEP 180  
 QY 181 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 240  
 DB 181 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 240  
 QY 241 KMWIPLTKVEQVEYLLRLD 261  
 DB 241 KMWIPLTKVEQVEYLLRLD 261

RESULT 3  
 US-10-425-114-48703  
 ; Sequence 48703, Application US/10425114  
 ; Publication No. US20040034898A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 48703  
 ; LENGTH: 297  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 700677130\_FLI.pep  
 US-10-425-114-48703

Query Match 100.0%; Score 1339; DB 12; Length 297;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-125;  
 Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAKAAEQSPDSGNVYTLASVREDLVQEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCG 60  
 DB 37 MAKAAEQSPDSGNVYTLASVREDLVQEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCG 96  
 QY 61 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHFGAASININKSI 120  
 DB 97 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHFGAASININKSI 156  
 QY 121 WKMYFKELLPLLATSGDGNVYTAANDLSLQSSRIHYGKFAVVKFRDAPQDYEP 180  
 DB 157 WKMYFKELLPLLATSGDGNVYTAANDLSLQSSRIHYGKFAVVKFRDAPQDYEP 216  
 QY 181 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 240  
 DB 217 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 276  
 QY 241 KMWIPLTKVEQVEYLLRLD 261  
 DB 277 KMWIPLTKVEQVEYLLRLD 297

RESULT 4  
 US-10-424-599-235878  
 ; Sequence 235878, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J  
 ; APPLICANT: Kovalic, David K  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 235878  
 ; LENGTH: 261  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_55025C.1.pep  
 US-10-424-599-235878

Query Match 98.2%; Score 1315; DB 12; Length 261;  
 Best Local Similarity 97.7%; Pred. No. 3.1e-123;  
 Matches 255; Conservative 6; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAKAAEQSPDSGNVYTLASVREDLVQEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCG 60  
 DB 1 MAKAAEQSPDSGNVYTLASVREDLVQEDTIIYGLIERAKFPSNSHTYDEKYAQIQGFCG 60  
 QY 61 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHFGAASININKSI 120  
 DB 61 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHFGAASININKSI 120  
 QY 121 WKMYFKELLPLLATSGDGNVYTAANDLSLQSSRIHYGKFAVVKFRDAPQDYEP 180  
 DB 121 WKMYFKELLPLLATSGDGNVYTAANDLSLQSSRIHYGKFAVVKFRDAPQDYEP 180



QY 181 IRAKDEGLMKLLTFTSVETVRKEKKAQVFGQVNLNSDDNDNENRKFDPSPVASSLY 240  
Db 181 IRAKDEGLMKLLTFTSVETVRKEKKAQVFGQVNLNSDDNDNENRKFDPSPVASSLY 240  
QY 241 KQWVPLTKVQVVEYLLRLD 261  
Db 241 KQWVPLTKVQVVEYLLRLD 261  
RESULT 5  
US-10-425-114-43606  
; Sequence 43606, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 43606  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3050-017-E9\_FLI.pgp  
US-10-425-114-43606

Query Match 72.4%; Score 970; DB 12; Length 199;  
Best Local Similarity 94.0%; Pred. No. 8.2e-89;  
Matches 187; Conservative 8; Mismatches 4; Indels 0; Gaps 0;  
QY 63 VEFVKNTEAIOAKAGRYKNPENAFPPENLPPSIVPSYFKQFLHFGAASININKSIWK 122  
Db 1 VEFVKNTEAIOAKAGRYKNPENAFPPENLPPSIVPSYFKQFLHFGAASININKSIWK 60  
QY 123 MYFKEKLLPLATSGDDGNYAQTAAANDLSLQSTSRRIHYGKFAVFKFDPADQDYELIR 182  
Db 61 MYFKEKLLPLATSGDDGNYAQTAAANDLSLQSTSRRIHYGKFAVFKFDPADQDYELIR 120  
QY 183 AKDKEGLMKLLTFTSVETVRKEKKAQVFGQVNLNSDDNDNENRKFDPSPVASSLYKN 242  
Db 121 AKDKEGLMKLLTFTSVETVRKEKKAQVFGQVNLNSDDNDNENRKFDPSPVASSLYKN 180  
QY 243 WVIPLTKVQVVEYLLRLD 261  
Db 181 WVIPLTKVQVVEYLLRLD 199

RESULT 6  
US-09-454-279-22  
; Sequence 22, Application US/09454279  
; Publication No. US20020184658A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; CURRENT APPLICATION NUMBER: US/09/454, 279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 22

; LENGTH: 265  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-454-279-22  
Query Match 61.1%; Score 818; DB 9; Length 265;  
Best Local Similarity 61.1%; Pred. No. 2.1e-73;  
Matches 162; Conservative 31; Mismatches 68; Indels 4; Gaps 2;  
QY 1 MAKAAEQSPDSG--NVYTLASVREDLVROEDTIYGLIERAKFPNSHNTYDEKYAQIQGF 58  
Db 1 MARVFESDSGSGCSNVLSLDLIESLIRQEDTIYGLIERAKFPNSPFAFESRCLDSGS 60  
QY 59 CGSLVEFVKNTEAIOAKAGRYKNPENAFPPENLPPSIVPSYFKQFLHFGAASININK 118  
Db 61 FSLTEFFVRETEIIQAKGRYEPENPFLENIPHSVPTHKYPSPALHPKALSININK 120  
QY 119 SIWKMYFKEKLLPLATSGDDGNYAQTAAANDLSLQSTSRRIHYGKFAVFKFDPADQDY 178  
Db 121 QIWDIYFKEKLLPLFVKPGDDGNYPTAASDLACLQALSRIHYGKFAVFKFDPADQDY 180  
QY 179 PLIRAKDEGLMKLLTFTSVETVRKEKKAQVFGQVNLNSDDNDNENRKF--DPSVA 236  
Db 181 PAIRAQDREALMKLLTTFEKVEEMVKRVQKKAETFGQEVKNSGYGDESKKXKVDPLLA 240  
QY 237 SSLYKNWVPLTKVQVVEYLLRLD 261  
Db 241 SRIYGEWVPLTKVQVVEYLLRLD 265

RESULT 7  
US-10-624-061-22  
; Sequence 22, Application US/10624061  
; Publication No. US20040019929A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/10/624,061  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: US/09/454,279  
; PRIOR FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 22  
; LENGTH: 265  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-624-061-22

Query Match 61.1%; Score 818; DB 15; Length 265;  
Best Local Similarity 61.1%; Pred. No. 2.1e-73;  
Matches 162; Conservative 31; Mismatches 68; Indels 4; Gaps 2;  
QY 1 MAKAAEQSPDSG--NVYTLASVREDLVROEDTIYGLIERAKFPNSHNTYDEKYAQIQGF 58  
Db 1 MARVFESDSGSGCSNVLSLDLIESLIRQEDTIYGLIERAKFPNSPFAFESRCLDSGS 60  
QY 59 CGSLVEFVKNTEAIOAKAGRYKNPENAFPPENLPPSIVPSYFKQFLHFGAASININK 118  
Db 61 FSLTEFFVRETEIIQAKGRYEPENPFLENIPHSVPTHKYPSPALHPKALSININK 120  
QY 119 SIWKMYFKEKLLPLATSGDDGNYAQTAAANDLSLQSTSRRIHYGKFAVFKFDPADQDY 178  
Db 121 QIWDIYFKEKLLPLFVKPGDDGNYPTAASDLACLQALSRIHYGKFAVFKFDPADQDY 180  
QY 179 PLIRAKDEGLMKLLTFTSVETVRKEKKAQVFGQVNLNSDDNDNENRKF--DPSVA 236  
Db 181 PAIRAQDREALMKLLTTFEKVEEMVKRVQKKAETFGQEVKNSGYGDESKKXKVDPLLA 240

```

QY 237 SSLYKXNVIPLTKEVQVEYLLRLD 261
DB 241 SRIYGEWLIPLTKEVQVEYLLRLD 265

RESULT 8
US-09-454-279-21
; Sequence 21, Application US/09454279
; Publication No. US20020184659A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: B1299 US NA
; CURRENT APPLICATION NUMBER: US/09/454,279
; PRIOR FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-454-279-21

```

```

Query Match 52.0%; Score 696; DB 9; Length 334;
Best Local Similarity 50.4%; Pred. No. 5.1e-61;
Matches 131; Conservative 46; Mismatches 75; Indels 8; Gaps 1;

QY 10 DSGNVYTLASVREDLVROEDTIYGLIERAKPPSNHTYDEKVAIQGFCGSLVEFVVK 69
DB 75 DESESLTEGIRNSLIQEDSIIFGLERAKYCYNADTYDPTAFMDGFGSLVEYVWG 134
QY 70 TEATQAKAGRYKNPEENAFPPENLPSPVPSYFQFLHPGAASININKSIWKYFKELL 129
DB 135 TEKLHAKVGRFKSPDEHPFDDLPPEMLPPLQYPKVLHFAADSININKIWNMYFRDLV 194
QY 130 PLLATSGDDGNYAQTAAANDLSLQSIIRIHYGKFAEVKFRDAPQDYEPILIRAKDKGL 189
DB 195 PRLVKGGDDGNYGSTAVCDIAICQLCKRIHYGKFAEAKFQASPEAYESAIAKQDKRL 254
QY 190 MKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENR-----KFDPSVASSLYK 241
DB 255 MDMLTFTPTVEDAIKKRVEMKTRTYGOEVKVGMEKEEEEEEGNESHVYKISPILVGDLG 314
QY 242 NWVPLTKVQVEYLLRLD 261
DB 315 DWIMPLTKVQVEYLLRLD 334

```

```

RESULT 9
US-10-624-061-21
; Sequence 21, Application US/10624061
; Publication No. US20040019929A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: B1299 US NA
; CURRENT APPLICATION NUMBER: US/10/624,061
; PRIOR FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: US/09/454,279
; PRIOR FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 334

```

```

; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-624-061-21

```

```

Query Match 52.0%; Score 696; DB 15; Length 334;
Best Local Similarity 50.4%; Pred. No. 5.1e-61;
Matches 131; Conservative 46; Mismatches 75; Indels 8; Gaps 1;

QY 10 DSGNVYTLASVREDLVROEDTIYGLIERAKPPSNHTYDEKVAIQGFCGSLVEFVVK 69
DB 75 DESESLTEGIRNSLIQEDSIIFGLERAKYCYNADTYDPTAFMDGFGSLVEYVWG 134
QY 70 TEATQAKAGRYKNPEENAFPPENLPSPVPSYFQFLHPGAASININKSIWKYFKELL 129
DB 135 TEKLHAKVGRFKSPDEHPFDDLPPEMLPPLQYPKVLHFAADSININKIWNMYFRDLV 194
QY 130 PLLATSGDDGNYAQTAAANDLSLQSIIRIHYGKFAEVKFRDAPQDYEPILIRAKDKGL 189
DB 195 PRLVKGGDDGNYGSTAVCDIAICQLCKRIHYGKFAEAKFQASPEAYESAIAKQDKRL 254
QY 190 MKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENR-----KFDPSVASSLYK 241
DB 255 MDMLTFTPTVEDAIKKRVEMKTRTYGOEVKVGMEKEEEEEEGNESHVYKISPILVGDLG 314
QY 242 NWVPLTKVQVEYLLRLD 261
DB 315 DWIMPLTKVQVEYLLRLD 334

```

```

RESULT 10
US-10-437-963-196439
; Sequence 196439, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 196439
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_92290C.1.pep
US-10-437-963-196439

```

```

Query Match 51.2%; Score 685.5; DB 16; Length 313;
Best Local Similarity 51.5%; Pred. No. 5.2e-60;
Matches 135; Conservative 45; Mismatches 8; Indels 1; Gaps 1;

QY 1 MAKAAEQSPDSGNVYTLASVREDLVROEDTIYGLIERAKPPSNHTYDEKVAIQGFCG 60
DB 52 VTKEEKQRIDQSEILTLNIRTSILVRQEDSIIFSLERAAQFCYNADIYKNAPHDVDFG 111
QY 61 SLVEFVVKNTAEIQAKAGRYKNPEENAFPPENLPSPVPSYFQFLHPGAASININKSI 120
DB 112 SLVEFVVKNTAEIQAKAGRYKNPEENAFPPENLPSPVPSYFQFLHPGAASININKSI 171
QY 121 WKMYFKELLPLIATSGDDGNYAQTAAANDLSLQSIIRIHYGKFAEVKFRDAPQDYEP 180
DB 172 WKMYFKELLPLIATSGDDGNYAQTAAANDLSLQSIIRIHYGKFAEVKFRDAPQDYEP 231
QY 181 IRAKQEGMLKLLTTSVETVRKVEKAVVFGQEVNLSDDNDNENR-KFDPSVASSL 239

```

```
Db 232 ITAQDCDQMLHLLTYETVERAIEHRVAKAFQGVDELGAEDNGAPPYKIRPSLVAEI 291
QY 240 YKNWVPLTKVEQVVEYLLRLD 261
Db 292 YSYRIMPLTKVEQVAYLLRLD 313

RESULT 11
US-09-454-279-14
; Sequence 14, Application US/09454279
; Publication No. US20020184658A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/09/454,279
; CURRENT FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; PRIOR FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-454-279-14

Query Match 50.9%; Score 681.5; DB 9; Length 257;
Best Local Similarity 52.1%; Pred. No. 1e-59;
Matches 134; Conservative 44; Mismatches 78; Indels 1; Gaps 1;

QY 6 EQSPDGNVYTLASVREDLVROEDTIYGLIERAKFPNSHTYDEKYAQIQGFCGSLVEF 65
Db 1 KORIDQSEILTDNIRTSVLRQEDSIIFSLLERAQFCYNADIYDKNAFHVDFGSLVEF 60
QY 66 VVKNTAIOAKAGRYKNPEENAFPPENLPPSVPSYFKQFLHPGAASININKSIWKYF 125
Db 61 MVRETEKLHQVGRYKSPDEHFFEDLPPLLPLOYKPKVLHPDIADSIINKEIKWYF 120
QY 126 KELLPLLATSGDGNVAQTAANDLSLLQSISSRIHYGKFAVAKFRDAPQDYEPILRAKD 185
Db 121 DELLPLRVKESDGNYSALCTICLQALSRIHYGKFAVAKFQESPEAYMPAIIAQD 180
QY 186 KEGMLKLLTFTSVETVRKVEKAVVFGQVNLNSDDNDNENR-KFDPSPVASSLYKNWV 244
Db 181 CDQLMHLTYETVERAIEHRVAKAFQGVDELGAEDNGAPPYKIRPSLVAEIYSYRI 240
QY 245 IPLTKEVQVVEYLLRLD 261
Db 241 MPLTKEVQVAYLLRLD 257

RESULT 12
US-10-624-061-14
; Sequence 14, Application US/10624061
; Publication No. US20040019929A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/10/624,061
; CURRENT FILING DATE: 2003-07-21
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/110,845
; PRIOR FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
```

```
; SEQ ID NO 14
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-624-061-14

Query Match 50.9%; Score 681.5; DB 15; Length 257;
Best Local Similarity 52.1%; Pred. No. 1e-59;
Matches 134; Conservative 44; Mismatches 78; Indels 1; Gaps 1;

QY 6 EQSPDGNVYTLASVREDLVROEDTIYGLIERAKFPNSHTYDEKYAQIQGFCGSLVEF 65
Db 1 KORIDQSEILTDNIRTSVLRQEDSIIFSLLERAQFCYNADIYDKNAFHVDFGSLVEF 60
QY 66 VVKNTAIOAKAGRYKNPEENAFPPENLPPSVPSYFKQFLHPGAASININKSIWKYF 125
Db 61 MVRETEKLHQVGRYKSPDEHFFEDLPPLLPLOYKPKVLHPDIADSIINKEIKWYF 120
QY 126 KELLPLLATSGDGNVAQTAANDLSLLQSISSRIHYGKFAVAKFRDAPQDYEPILRAKD 185
Db 121 DELLPLRVKESDGNYSALCTICLQALSRIHYGKFAVAKFQESPEAYMPAIIAQD 180
QY 186 KEGMLKLLTFTSVETVRKVEKAVVFGQVNLNSDDNDNENR-KFDPSPVASSLYKNWV 244
Db 181 CDQLMHLTYETVERAIEHRVAKAFQGVDELGAEDNGAPPYKIRPSLVAEIYSYRI 240
QY 245 IPLTKEVQVVEYLLRLD 261
Db 241 MPLTKEVQVAYLLRLD 257

RESULT 13
US-10-437-963-151855
; Sequence 151855, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 151855
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_51960C.1.pep
US-10-437-963-151855

Query Match 50.7%; Score 678.5; DB 16; Length 255;
Best Local Similarity 53.1%; Pred. No. 2e-59;
Matches 135; Conservative 43; Mismatches 63; Indels 13; Gaps 3;

QY 16 TLASVREDLVROEDTIYGLIERAKFPNSHTYDEKYAQIQGFCG--SLVEFVVKNTTEAI 73
Db 7 SLAAVRDALVREDSIVFALIERARRPRNAPAYAAAAA-----AGGRSLAEFFVREAEVL 61
QY 74 QAKAGRYKNPEENAFPPENLPPSVPSYFKQFLHPGAASININKSIWKYFVKELLPLLA 133
Db 62 HAKAGYQKQKEDVFFFPQDLPSPFLPTKYPKVLHSPASVSVNDALWKYFNELLPLFT 121
QY 134 TSGDGNVAQTAANDLSLLQSISSRIHYGKFAVAKFRDAPQDYEPILRAKDKEGLMKLL 193
Db 122 VDGDDGNVAETVALDFACLKALSRRIHYGKFAVAKFRDAPQDYEPILRAKDKALMNL 181
```

QY 194 TFTSVEETVRKVRKAVVFGQEVNLSDDNDNENR-----KTPSPVASSLYKNWVPL 247  
Db 182 TFKAEEVKRVRKVRKARIFQNTVLEDNADKQEGNAGDSECKYNPEVLSKLYDLWVPL 241  
QY 248 TKEVQVEYLLRLD 261  
Db 242 TKDVEVEYLLRLD 255

RESULT 14  
US-09-454-279-6  
; Sequence 6, Application US/09454279  
; Publication No. US20020184658A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Pamodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 132  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-454-279-6

Query Match 50.3%; Score 673; DB 9; Length 132;  
Best Local Similarity 98.5%; Pred. No. 2.7e-59;  
Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 10 DSGNVYTLASVREDLVROEDTIIYGLIERAKFPNSHTYDEKYAQIQGFCGSLVEFVVK 69  
Db 1 DSGNVYTLASVREDLVROEDTIIYGLIERAKFPNSHTYDEKYAQIQGFCGSLVEFVVK 60  
QY 70 TEAIQAKAGRYKNPEENAFFPENLPPSIVPSYSPKQFLHPGAASININKSIWKMYFKELL 129  
Db 61 TEAIQAKAGRYKNPEENAFFPENLPPSIVPSYSPKQFLHPGAASININKSIWKMYFKELL 120  
QY 130 PLLATSGDDGNY 141  
Db 121 PLLATSGDDGNY 132

RESULT 15  
US-10-624-061-6  
; Sequence 6, Application US/10624061  
; Publication No. US2004001929A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Pamodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/10/624,061  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: US/09/454,279  
; PRIOR FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 132  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-10-624-061-6

Query Match 50.3%; Score 673; DB 15; Length 132;  
Best Local Similarity 98.5%; Pred. No. 2.7e-59;  
Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 10 DSGNVYTLASVREDLVROEDTIIYGLIERAKFPNSHTYDEKYAQIQGFCGSLVEFVVK 69  
Db 1 DSGNVYTLASVREDLVROEDTIIYGLIERAKFPNSHTYDEKYAQIQGFCGSLVEFVVK 60  
QY 70 TEAIQAKAGRYKNPEENAFFPENLPPSIVPSYSPKQFLHPGAASININKSIWKMYFKELL 129  
Db 61 TEAIQAKAGRYKNPEENAFFPENLPPSIVPSYSPKQFLHPGAASININKSIWKMYFKELL 120  
QY 130 PLLATSGDDGNY 141  
Db 121 PLLATSGDDGNY 132

Search completed: September 25, 2004, 03:00:23  
Job time: 74 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2004, 02:42:06 ; Search time 24 Seconds  
(without alignments)  
1046.083 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKAAEQSPDSGNVYTLASV.....NWVPLTKEVQVEYLLRLD 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: Pir1.\*

2: Pir2.\*

3: Pir3.\*

4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	818	61.1	265	T50796	chorismate mutase
2	696	52.0	334	S38958	chorismate mutase
3	498.5	37.2	256	A45921	chorismate mutase
4	462.5	34.5	251	T37784	probable chorismat
5	283	21.1	82	T14902	chorismate mutase
6	188	14.0	70	T14901	chorismate mutase
7	105	7.8	360	G95041	hypothetical prote
8	103	7.7	781	B64135	ribonucleoside-dip
9	95.5	7.1	587	F90044	hypothetical prote
10	95.5	7.1	1556	S59393	probable membrane
11	93.5	7.0	295	S56296	probable membrane
12	93.5	7.0	1134	D75014	hypothetical prote
13	92	6.9	371	A71683	hypothetical prote
14	92	6.9	605	S64410	probable membrane
15	92	6.9	823	AF3361	endopeptidase Ia (
16	91.5	6.8	631	AF0932	thiamin biosynthes
17	91	6.8	439	H81414	probable heat choc
18	90.5	6.8	554	S37936	regulatory protein
19	90	6.7	946	JC5575	inter-alpha-trypsi
20	90	6.7	1015	JC6552	DNA topoisomerase
21	89.5	6.7	21	T70163	hypothetical prote
22	89.5	6.7	1187	C84568	conserved hypothet
23	89	6.6	431	H70302	DNA topoisomerase
24	88.5	6.6	668	G90241	hypothetical prote
25	88.5	6.6	578	T44444	probable homeodoma
26	88.5	6.6	680	H84774	hypothetical prote
27	88.5	6.6	1468	F70466	RNA polymerase bet
28	88.5	6.6	1944	A59438	KIAA1474 protein [
29	88	6.6	1944	F71061	hypothetical prote

#### RESULT 2

S38958

Chorismate mutase (EC 5.4.99.5) precursor - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 07-Mar-2003

#### ALIGNMENTS

##### RESULT 1

T50796

Chorismate mutase CM2 - Arabidopsis thaliana

N:Alternate names: protein T30N20\_140

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 07-Mar-2003

C:Accession: T50796

R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew

submitted to the Protein Sequence Database, July 2000

A:Reference number: 225240

A:Accession: T50796

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-265 <BEV>

A:Cross-references: EMBL:AL365234

A:Experimental source: cultivar Columbia; BAC clone T30N20

C:Genetics:

A:Map position: 5

A:Introns: 78/3; 107/3; 155/3; 186/3

A>Note: T30N20\_140

C:Superfamily: Chorismate mutase of the AroQ class, eukaryotic type

Query Match	61.1%;	Score 818;	DB 2;	Length 265;
Best Local Similarity	61.1%;	Pred. No. 9e-59;		
Matches	162;	Conservative	31;	Mismatches 68; Indels 4; Gaps 2;
QY	1	MAKAAEQSPDSG--NVYTLASVREDLVQRQEDTIYGLIERAKPPSNHGTDEKTAQIQGF	58	
DB	1	MARVFESDGSQCSNVLSDLIRESLRQEDTIVFSLIERAKFPPLNSPAFESRCLDSGS	60	
QY	59	CGSLVEFVVKNTAIOAKAGYKKNPEANAFPPENLPSPSIVPSYFKQFLHFGAASINIK	118	
DB	61	FSSUTEFVRTEIIOAKVGRIYEPENPFLENIPHSVFTTKYKPSALHPKALSVNINK	120	
QY	119	SIWKMYFKELLPLLATSGDGNVAQTAANDLSLQISRIIRIHGKFAVVKFRDAPQDYE	178	
DB	121	QIWIDYFKELLPLFKPGDDGNYPSTAASDLACIQLASRIIRIHGKFAVVKFRDAPQDYE	180	
QY	179	PLIAKDKGKMLLTFTSVEETVRKVEKAVVFGQEVNLSDDNENRKF--DPSVA	236	
DB	181	PAIRAQREALMKLLTFKVEEMVKRVQKKAETFGQEVFNFGSGYDSEKKYKVDPLLA	240	
QY	237	SSLYKNWVPLTKVEQVEYLLRLD	261	
DB	241	SRIYGEWLLPLTKLVEVEYLLRLD	265	

30	88	6.6	495	2	T33065	hypothetical prote
31	88	6.6	655	1	S50096	probable splicing
32	88	6.6	737	2	S63453	probable RNA helic
33	87.5	6.5	596	2	E96335	FUSION, methionine
34	87.5	6.5	2829	2	A42771	reticulocyte-bind
35	87	6.5	452	2	A83758	aldehyde dehydrog
36	87	6.5	701	2	T17243	hypothetical prote
37	87	6.5	936	2	A36384	probable membrane
38	86.5	6.5	1653	2	A36349	clathrin heavy cha
39	86	6.4	265	2	D86715	conserved hypothet
40	86	6.4	428	2	AC1393	homoserine dehydro
41	86	6.4	526	1	T45058	phosphoprotein pho
42	86	6.4	638	2	A82488	hypothetical prote
43	86	6.4	843	2	H90514	lipoprotein (impor
44	86	6.4	946	2	S54354	inter-alpha-inhibi
45	85.5	6.4	918	2	S45872	hypothetical prote



## RESULT 5

T14902  
Chorismate mutase (EC 5.4.99.5) 2, cytosolic - parsley (fragment)  
C:Species: Petroselinum crispum (parsley)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 07-Mar-2003  
C:Accession: T14902  
R:Batiz, O.; Logemann, E.; Reinold, S.; Hahlbrock, K.  
submitted to the EMBL Data Library, July 1997  
A:Description: Extensive reprogramming of cellular metabolism by fungal elicitor or infection  
A:Reference number: Z18251  
A:Accession: T14902  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-82 <BAT>  
A:Cross-references: EMBL:AF012867; NID:g2352930; PID:g2352931  
A:Gene: CCM2  
C:Function:  
A:Description: intramolecular transferase; isomerase  
A:Pathway: aromatic amino acid biosynthesis  
C:Superfamily: chorismate mutase of the AroQ class, eukaryotic type  
C:Keywords: aromatic amino acid biosynthesis; intramolecular transferase; isomerase; shi

Query Match 21.1%; Score 283; DB 2; Length 82;  
Best Local Similarity 63.4%; Pred. No. 3.4e-16;  
Matches 52; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 88 FFPENLPSIVSPYSPKQFLHPGAASININIKWYFKELLPLLATSGDGNVYQAQTAAN 147

Db 1 FPDNLPSLLQYNHPQLLHTPTAAINININWDAVYNQLPLTTTGGDGNVYPTATS 60

QY 148 DLSLQISRRRIHYGKFAEVK 169

Db 61 DLQCLQASRRRIHYGKFAEAK 82

## RESULT 6

T14901  
Chorismate mutase (EC 5.4.99.5) 1, chloroplast - parsley (fragment)  
C:Species: Petroselinum crispum (parsley)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 07-Mar-2003  
C:Accession: T14901  
R:Batiz, O.; Logemann, E.; Reinold, S.; Hahlbrock, K.  
submitted to the EMBL Data Library, July 1997  
A:Description: Extensive reprogramming of cellular metabolism by fungal elicitor or infection  
A:Reference number: Z18251  
A:Accession: T14901  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-70 <BAT>  
A:Cross-references: EMBL:AF012866; NID:g2352928; PID:g2352929  
C:Genetics:  
A:Gene: pCM1  
A:Genome: nuclear  
C:Function:  
A:Description: intramolecular transferase; isomerase  
A:Pathway: aromatic amino acid biosynthesis  
C:Superfamily: chorismate mutase of the AroQ class, eukaryotic type  
C:Keywords: aromatic amino acid biosynthesis; chloroplast; intramolecular transferase; i

Query Match 14.0%; Score 188; DB 2; Length 70;  
Best Local Similarity 51.4%; Pred. No. 1.3e-08;  
Matches 36; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

QY 90 PENLPSIVSPYSPKQFLHPGAASININIKWYFKELLPLLATSGDGNVYQAQTAANDL 149

Db 1 PEHLPEIPGPLOYPQVLPVHTESININIKWYFNRLVPLRVKEDGNGYATAVCDT 60

QY 150 SLIQSISRRI 159

Db 61 ICLOALSKRI 70

## RESULT 7

G95041  
hypothetical protein SP0355 [imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
C:Accession: G95041  
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heide-  
son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel, I.  
on, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: G95041  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <KUR>  
A:Cross-references: GB:AE005672; PIDN:AAK74528.1; PID:g14971828; GSPDB:GN00164; TIGR:SP48  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP0355

Query Match 7.8%; Score 105; DB 2; Length 360;

Best Local Similarity 24.7%; Pred. No. 0.64;  
Matches 55; Conservative 34; Mismatches 76; Indels 58; Gaps 11;

QY 31 IYGLIERAKF-----PSN-SHTYDEKYAQIQGFGSLVFVVVXNTEATQAKAGRYKNP 83

Db 26 IVYDLLESKGVESFYNPSDFQYLYKEVRQKQ-----SFTKKQADAILYIPGGYFGE 78

QY 84 EENAFPPENLPPSIVSPYSPKQFLHPGAASININIKWYFKELLPLLATSGDGNVYQA 143

Db 79 GHNARFRDNL-----IQKRFPLIG-----IWASYFKKPIGVIGIGAGPN---- 118

QY 144 TAANDLSLQISRRRIHYGKFAEVKFRD-----AP-QDYEPILIRAKDEGLM 190

Db 119 ---NDSLMNYGKRIINHAQFIT-VRDRESFDSLKLHSPSPVHETFDLISSKLREKT 174

QY 191 KLLFTTSVEETVRKV-----EKKAV-VFGQEVNLSDDNDN 226

Db 175 EQCLQKREARKDKIILVHYHNSKKALEKFAESISLPLENNP 217

## RESULT 8

B64135  
ribonucleoside-diphosphate reductase (EC 1.17.4.1) alpha chain - Haemophilus influenzae  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 18-Jun-1999  
C:Accession: B64135  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
; D.M.; Brandon, R.C.; Pine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: B64135  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-781 <TIGR>  
A:Cross-references: GB:U32839; GB:I42023; NID:g1574508; PIDN:AA23305.1; PID:g1574509; TJ  
C:Genetics:  
A:Start codon: GTG  
C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase large chain  
C:Keywords: DNA replication; iron-sulfur protein; oxidoreductase

Query Match 7.7%; Score 103; DB 2; Length 781;

Best Local Similarity 22.9%; Pred. No. 2.7;  
Matches 53; Conservative 36; Mismatches 104; Indels 38; Gaps 9;

QY 49 DEKVAQIQGFGSLVFVWK-----NTEATQAKAGRYKNPEENAFPPENLPPSIVSPYS 102

Db 256 DSDLSINATASAIKYYVSQAGIGINAGATGALGSEIRGGE--AFH-----TGCIFPK 308  
QY 103 FKQF-----LHPGAASININKSIWMYFKELLPLLATSGDDGNYACTAANDLSLLQ 153  
Db 309 YFQTAVKSCSGVVGGAATLY--PIWLEAENLLVLKNNRGVEDNRVHMDYGVQLNK 366  
QY 154 SISRIHVGKFAEVKFRDADQDVEPLIRAKDKEGLMKLLTFTSVEETVRKRVKKAVWF 213  
Db 367 LMVQLIKGSEITLSPSDVFGLEZAFFADQDK--FEELYVKYEQDPTIRKRTYKAVEIF 424  
QY 214 G---QE-----VNLSDNDNENRKFDPSPVASSLYKNWVPL--TKEVQ 252  
Db 425 SLLMQERASTGRHYIQNVHDCNTHSPDPQVAPVRQSNLCLEIALPKPLQ 475

RESULT 9  
F90044  
Hypothetical protein SA2217 [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: F90044  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: F90044  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-587 <KUR>  
A:Cross-references: GS:BA000018; PID:g13702378; PIDN:BA843519.1; GSPDB:GN00149  
A:Experimental source: Strain N315  
C:Genetics:  
A:Gene: SA2217  
C:Superfamily: beta-1,2-glucan export protein chvA; ATP-binding cassette homology

Query Match 7.1%; Score 95.5; DB 2; Length 587;  
Best Local Similarity 22.9%; Pred. No. 7.4;  
Matches 56; Conservative 45; Mismatches 101; Indels 43; Gaps 10;  
QY 45 SHTYDEKIAQICGFCGSIVP-----VVKNTAQAQAGRYKNPEEN-----AFPE-- 91  
Db 186 NETVAEQMKSNMNSAIVEFIEGIEVIKTFNQSCSSYKYYKDAVDNKKIHTLWFKNTW 245  
QY 92 --NLPPSIVPSYSPKQFLHPGAASININKSIWMYFKEL--LPLLATSGDDGNYAQT 145  
Db 246 GYNLGSVLP--TFLGTLPGVMLISINQLNYAEFFLCIVLSGVVAPIKNPTNYVNL 304  
QY 146 ANDLSLLQSIRRIHYGKFAEVKFRDADQDVEPLIR-----AKDKEGLM-KLLTPT-- 196  
Db 305 KSIQVALTENVQILSLEBLVLTSPK-KPQRYEAFNNGVPSYKNDKDLVFKHLSPTVP 363  
QY 197 -----SVETVRKVEKAVVFGQEVNLS-DNDNENRKFDPSPVASSLYKNW 243  
Db 364 ENNFAIVGASGSGKSTAKLISRWDTSGTEITIGGINIKDLESQNDLVGFGVDNF 423  
QY 244 VIPLT 248  
Db 424 LLNLT 428

RESULT 10  
S59393  
probable membrane protein YLR247c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein L9672.14  
C:Species: Saccharomyces cerevisiae  
C:Date: 30-Nov-1995 #sequence\_revision 23-Feb-1996 #text\_change 19-Apr-2002  
C:Accession: S59393  
R:Johnson, D.  
A:Description: The sequence of S. cerevisiae cosmid 9672.

A:Reference number: S59386  
A:Accession: S59393  
A:Molecule type: DNA  
A:Residues: 1-1556 <JOH>  
A:Cross-references: EMBL:U20865; NID:G662330; PID:G662344; GSPDB:GN00012; MIPS:YLR247c  
A:Experimental source: strain S288C (AB972)  
C:Genetics:  
A:Gene: MIPS:YLR247c  
A:Cross-references: SGD:S0004237  
A:Superfamily: RING finger homology  
C:Superfamily: RING finger homology  
C:Keywords: transmembrane protein  
F:1235-1282/Domain: RING finger homology <RRN>  
F:1236-1252/Domain: transmembrane #status predicted <TMM>

Query Match 7.1%; Score 95.5; DB 2; Length 1556;  
Best Local Similarity 21.3%; Pred. No. 28;  
Matches 65; Conservative 46; Mismatches 79; Indels 115; Gaps 16;  
QY 3 KAAEQSPDSGNVYTLASVREDLVROEDTIYGLIERAKFPNSHTYDEKIAQI----- 55  
Db 1307 EGADSSQNSN-----ENSIINXSEVEKLFQNKY---EQFHINEVHOIH 1349  
QY 56 --QFGCSIVPVPKNTAIAQAGRYKNPEENAFPPENLPPSIVPSYSPKQFLHPGAAS 113  
Db 1350 IKESP-GAKIDFVILISYL-----RLKSEQENAD-----PPQVI-LYSQK--TEY 1391  
QY 114 ININKSIWMYFKELLPLLATSGDDG---NVAQTAA----- 146  
Db 1392 LKVIQKVLKYLHIEHLACLSNTANVGETINFPKQPSVTCLLNVTGLAGNLINAKHI 1451  
QY 147 -----NDLSLLQSIRRIHYGKFAEVKFRDADQDVEP-----LIRAKDKEGLMKLLT 194  
Db 1452 FLDDPLNNSDELQAMGRNRIG-----QDETFWVNFMIKNTVEINILRYKC 1499  
QY 195 FTSVETVRKVEKAVVFGQEVNLSDDNDNE---NRKFDPSVASSLYKNWVPLTKEV 251  
Db 1500 I--LEERRKKEKSKK---GDKYEAQDETNEESDDAKFEISV-----VDOEV 1542  
QY 252 QVEYL 256  
Db 1543 SNEHL 1547

RESULT 11  
S56296  
probable membrane protein YFR041c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein F018  
C:Species: Saccharomyces cerevisiae  
C:Date: 02-Sep-1995 #sequence\_revision 19-Oct-1995 #text\_change 23-Sep-2002  
C:Accession: S56296; S62252; S63788  
R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasanun  
submitted to the EMBL Data Library, May 1995  
A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces ce  
A:Reference number: S56186  
A:Accession: S56296  
A:Molecule type: DNA  
A:Residues: 1-295 <NUR>  
A:Cross-references: EMBL:D50617; NID:G836685; PID:d1009921; PID:G836796; MIPS:YFR041c  
R:Murakami, Y.  
submitted to the EMBL Data Library, December 1994  
A:Reference number: S62230  
A:Accession: S62252  
A:Molecule type: DNA  
A:Residues: 1-295 <NUR>  
A:Cross-references: EMBL:D44597; NID:G871938; PID:d1008597; PID:G871940  
R:Xi, T.; Naitou, M.; Hagiwara, H.; Ozawa, M.; Sasanuma, S.I.; Sasanuma, M.; Tsuchiya, J  
yeast 12, 149-167, 1996  
A:Title: Analysis of a 36.2 kb DNA sequence including the right telomere of chromosome VI  
A:Reference number: S63787; MUID:96287652; PMID:8686379  
A:Accession: S63788  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA



A;Residues: 1-295 <EKI>  
A;Cross-references: EMBL:D44597; NID:g871938; PID:d1008597; PID:g871940  
C;Genetics:  
A;Cross-references: SGD:S0001937  
A;Map position: 6R  
A;Note: YPR041c  
C;Superfamily: yeast membrane protein YPR041c; dnaJ amino-terminal homology  
C;Keywords: transmembrane protein  
F;8-24/Domain: transmembrane #status predicted <TM1>  
F;44-108/Domain: dnaJ amino-terminal homology <DNJ>  
F;133-149/Domain: transmembrane #status predicted <TM2>

Query Match 7.08; Score 93.5; DB 2; Length 295;  
Best Local Similarity 22.1%; Pred.No.4.2;  
Matches 56; Conservative 42; Mismatches 96; Indels 59; Gaps 13;

Qy 15 YTTLASVEDLVQRDDTIYGLIERAKFPSSHTDEKYAQIQFGSLVFWKNTIAIQ 74  
Db :  
19 YAFTTIEIFQLQNEI-----STKYGPDWNFY--KFLKLPLQNSSTKEITQLRKLS 70  
Qy 75 AKGRYNKPENAFPE-NLPPTSIPYS----FKQFLHPCAASININKSIWKMYFKELL 129  
Db :  
71 KKYHPDKNPKRYKLYERLNLATQLSNSSNRKIYDYVLQNGFFNYDPHKG--GFYFSRMK 128  
Qy 130 P---LLATSGDDGNYAQTANDILLSQTSRRRIHYGVFAEVKFRDAPQDIYPLRAKD 185  
Db 129 PKTWFLAFTWIVVNICQYI---ISIIQYRSQRSRIENFISOCKQQD-----D 173  
Qy 186 KEGI-MKLLAFTSVEEVRKRVEKKAVVF-----GQEVNLNSDDNDNENRKFD 233  
Db 174 TNGLGVKQLTFKQHEKD-----EGKSLVRFSDVVVPDGGSETLISPDLTK-----P 222  
Qy 234 SVASSLYKNWVIP 246  
Db 223 SVKNCLF--WRIP 233

RESULT 12  
D75014  
Hypothetical protein PAB1257 - Pyrococcus abyssi (strain Orsay)  
C;Species: Pyrococcus abyssi  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
C;Accession: D75014  
R;anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru  
A;Reference number: A75001  
A;Accession: D75014  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1134 <XAW>  
A;Cross-references: GB:AJ248289; GB:AL096836; NID:g5458950; PIDN:CAB50554.1; PID:e151645  
A;Experimental source: strain Orsay  
C;Genetics:  
A;Gene: PAB1257

Query Match	7.0%;	Score	93.5;	DB	2;	Length	1134;
Best Local Similarity	22.8%;	Pred. No.	26;				
Matches	47;	Conservative	41;	Mismatches	71;	Indels	47;
Gaps	9;						

  

Qy	18	ASVREDLVRQEDNIIVGLIERAPPSNSHYDEKYAQIQGFCGSLVEFVYKNTAIIQAKA	77
Db	654	SEIKDLVRVEKISEIIII-----IYNOKYWCYKSRHASFV-----EKITGKA	697
Qy	78	GRYKXNPENAPFPENLPPSIVPSYFQF--LH-----PGAASININKSIWMYFKELL	129
Db	698	RVYKDECENA---EDRISNLLNLLGFPQYIMLHSMGMGIVEGVAEINIVSTIGBVSKE-L	753
Qy	130	PLLATSGDDGNVYAQTANDLSL-LQISRSRIHYKFFVAEVKFRDAPQDIEPLRAKDKG	188
Db	754	SRGTGNIIEVTKLONIRUGDDJISLEDIAKK-----SRDYNFEPLHSVEETI	798
Qy	189	LMKLLLTFTSV---EETVYKREVEKKAV	211

| : | : | : | : | : | : |

Dd       799 LRELLAYKGISTREHELKSEIDKLSV 824

RESULT 13  
A71683  
hypothetical protein RP278 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii  
C>Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
C:Cessionion: A71683  
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.  
Nature 396, 133-140, 1998  
A>Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A:Reference number: A71630; MUID:99039499; PMID:9823893  
A:Accession: A71683  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-371 <AND>  
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA14739.1; PID:g386083  
A:Experimental source: Strain Madrid E  
C:Genetics:  
A:Gene: RP278

Query Match	6.9%;	Score 92;	DB 2;	Length 371;	
Best Local Similarity	22.6%;	Pred. No. 7.5;			
Matches	64;	Conservative 34;	Mismatches 91;	Indels 94;	Gaps 14
Qy	14	VYTLASVREDLVROEDTIYGLI	-----ERAKF-----	41	
Db	117	LYSLTKETDLIKANNIPESLTPDDLRLLKGENLTPQEQEBERKKFEYLSILGSIIDDT	176		
Qy	42	-PSNHTTDEKVAQIQFCGSLV--EFVGNTEAIQAKAGRYKYNPENAFPENLPPSIV	98		
Db	177	KKSNEH-YDKRANEINEQNLNTIINEFRNLQOEKKIEGKKTQASEA-----LNKKUK	230		
Qy	99	PSYSKQF-----LHPGAASININKSIWRYFKELLPLLATSGDPNYAQTA	146		
Db	231	PIY--KOMDESREELFKLAELIPQYAQANIDKHA-KLYAKQYQTKIE--NDPNY-----	279		
Qy	147	NDLSSLQISRIIHVGKFAVKEFRDAPQDYELIRAKOKEGIMKLLITTSVEETVRKRV	206		
Db	280	KLEKLIQIVSKIEYSK---KSKIND-----IINPNSTVVKLLLEKIDEKISSKE	329		
Qy	207	EKKAV-VFGQEVNLNSDDNDNENKFDPSVASSLYKRWIPLT	248		
Db	330	MTEANLVKQEVNLT-----KPFITPLT	353		

RESULT 14  
S64410  
probable membrane protein YGR103w - Yeast (Saccharomyces cerevisiae)  
N;Alternate names: hypothetical protein G5933  
C;Species: Saccharomyces cerevisiae  
C;Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 19-Apr-2002  
C;Accession: S64410; S64408  
R;Wedler, H.; Schaeffe, M.; Wedler, E.; Wambutt, R.  
submitted to the Protein Sequence Database, May 1996  
A;Reference number: S64356  
A;Accession: S64410  
A;Molecule type: DNA  
A;Residues: 1-605 <WED>  
A;Cross-references: EMBL:Z72888; NID:g1323158; PID:g1323159; MIPS:YGR103w  
A;Experimental-source: strain S288C  
R;Hernandez, K.; Weber, N.; Wipili, P.; Schmidheini, T.  
submitted to the Protein Sequence Database, May 1996  
A;Reference number: S64392  
A;Accession: S64408  
A;Molecule type: DNA  
A;Residues: 1-18 <HER>  
A;Cross-references: EMBL:Z72888; MIPS:YGR103w  
A;Experimental source: strain S288C  
C;Genetics:  
A;Gene: SGD:RRP13

A:Cross-references: SGD:S0003335  
A:Map position: 7R  
C:Keywords: transmembrane protein  
F:217-233/Domain: transmembrane #status predicted <TMM>

Query Match 6.9%; Score 92; DB 2; Length 605;  
Best Local Similarity 20.5%; Pred. No. 15;  
Matches 62; Conservative 50; Mismatches 116; Indels 74; Gaps 13;

QY 2 AKAAQSPDSGNVYTLASVRE-----DLVRQEDTIYGLIERAKFPFSNSHTYDEKYAQ 54  
DB 102 AKRLEENRDS---YTLDHIIKERYSPFDAIRDIDDALNMLFLFSLNPSTNQVSSKIIND 158  
QY 55 IQGFCGSLVEFVVKV-----TEATQAKAGRYKNPEENAFPPENLPPSI-- 97  
DB 159 AQKICNQWLAYAKERLVKRVKSVKGVYQYQANIKGEVRLVFPK---FPENIPSDVDF 215  
QY 98 VPSYGFQKFLHPCAASINIKSIWYFKELLPLIATSGDDGNYAQTAANDLSLQISIR 157  
DB 216 RIMLTFLEFY-----STLHFVLYKLY-----TDSGLIYPPKLDLKKDKIIS- 257  
QY 158 RIHYGKFAVAKFRDAPDYEPL-----IRAKQKGLMKLLTF-----TSVEETVR 203  
DB 258 --GLSSYIIIESRQEDSLKLDPTETEEDVKVESLDASTKSAALNADEANTDETKEEEOE 315  
QY 204 KEVEK-KAVVFGQEVNLSDDNDNENR-----KPDPSVASSLYKNWVIPLTKVQVE 254  
DB 316 KQEKQKQEKQNEETELDTFEDNNKNGDILIQSKYDSPA-SLPSAFVYVVSREVPID 374  
QY 255 YL 256  
DB 375 IL 376

Search completed: September 25, 2004, 02:54:47  
Job time : 27 secs

QY 176 DYEPLIRAKDKEGLMKLLTF-----TSVEETVRKRVKKAIVFGQEVNLSN----- 220  
DB 202 KOEMLSVLSVRERLKAUSFMEABISVLQVEKRIRSRVRQMEKTQREYLYNEQMKAIQK 261  
QY 221 ----SDNDNENRKFDPSPVASSLYKNWVIPLTKVQVEYL--LRRL 260  
DB 262 ELGDSGDSGRDEVAIEERITKT-----KLSKEAREKALAEELKKL 300

RESULT 15  
AF3361  
endopeptidase La (EC 3.4.21.53) [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 02-Jun-2003  
C:Accession: AF3361  
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goltzman, E.; Selkov, E.; Eizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AF3361  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-823 <KUR>  
A:Cross-references: GB:AE008917; PIDN:AAL52057.1; PID:gl7982825; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0876  
A:Map position: I  
C:Superfamily: ATP-dependent Lon protease  
C:Keywords: hydrolase; serine proteinase

Query Match 6.9%; Score 92; DB 2; Length 823;  
Best Local Similarity 22.0%; Pred. No. 22;  
Matches 63; Conservative 44; Mismatches 85; Indels 94; Gaps 14;

QY 4 RAEQSPDSGNVY---TLASVREDLVRQEDTIYGLIE---RAKFP--SNSHTYDEKYAQI 55  
DB 80 AADDDPAPDAIYEIGTIANVLO-LKLPLDGVIVKVLVEGTARAKISKFTDREDYHEAYA-- 136  
QY 56 QGFCGSLVEFVVKVTEAIQAKAGRYKNPEENAFPPENLPPSIYPSYKQFLHPCAASIN 115  
DB 137 -----AALQ-----EPEDAVEIEALARSVVS--DPENY-----VK 165  
QY 116 INKSIWKYFKELLPLIATSGDDGNYAQTAANDLSLQISIRIHYGKFAVAKFRDAPQ 175  
DB 166 LNKKI-----SPEVVGTSQIDDYSKLADTVASHLAI-----KIPF 201

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2004, 01:26:35 ; Search time 18 seconds

(without alignments)  
755.017 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 NAKAAEQSPDSGNVYLASV.....NWVPLTKVEQVVEYLLRLD 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	698	52.1	334	1	CHMU_ARATH
2	498.5	37.2	256	1	CHMU_YEAST
3	462.5	34.5	251	1	CHMU_SCHPO
4	103	7.7	756	1	RIRI_HAEIN
5	101	7.5	737	1	SUV3_SACDO
6	94.5	7.1	761	1	RIRI_BUCBP
7	93.5	7.0	295	1	YFL1_YEAST
8	93	6.9	812	1	LON_BRUAB
9	92	6.9	605	1	PESC_YEAST
10	92	6.9	812	1	LON_BRUME
11	92	6.9	812	1	LON_BRUSU
12	91.5	6.8	510	1	ATPA_BUCAP
13	91.5	6.8	631	1	THIC_SALTI
14	91.5	6.8	832	1	ALP6_SCHPO
15	91	6.8	439	1	HSLU_CAMJE
16	90.5	6.8	554	1	HAP4_YEAST
17	90	6.7	693	1	EFG_STEMU
18	90	6.7	946	1	ITB2_MESAU
19	88.5	6.6	1468	1	RPOB_AQUAE
20	88	6.6	737	1	SUV3_YEAST
21	87.5	6.5	631	1	THIC_SALTY
22	87.5	6.5	789	1	CAD9_HUMAN
23	87.5	6.5	2869	1	RBPI_PLAVB
24	87	6.5	795	1	CDB5_HUMAN
25	87	6.5	936	1	YG2K_YEAST
26	86.5	6.5	1653	1	CLH_YEAST
27	86	6.4	946	1	ITB2_MOUSE
28	85.5	6.4	918	1	IMB2_YEAST
29	85	6.3	299	1	COMQ_BACSU
30	85	6.3	381	1	ARGJ_TETH
31	85	6.3	601	1	SG2_RANRI
32	84	6.3	1051	1	VP2_AHSV6
33	83.5	6.2	373	1	YIA6_YEAST

## RESULT 1

ID	CHMU_ARATH	STANDARD;	PRT;	334 AA.
AC	P42738; Q9LS75; Q9SUJ5;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Chorismate mutase, chloroplast precursor [EC 5.4.99.5] (CM-1).			
GN	CM1 OR A13G29200 OR M0211.4.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.			
OX	NCBI_taxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94039841; PubMed=8224252;			
RA	Eberhard J., Raesecke H.-R., Schmid J., Amrhein N.;			
RT	"Cloning and expression in yeast of a higher plant chorismate mutase.			
RT	Molecular cloning, sequencing of the cDNA and characterization of the			
RT	Arabidopsis thaliana enzyme expressed in yeast.";			
RL	FEBS Lett. 334:233-236(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Landsberg erecta;			
RA	Kuhn R., Vogt E., Schmid J., Amrhein N., Schaller A.;			
RT	"Expression analysis of Arabidopsis thaliana genes for plastidic (CM1)			
RT	and cytosolic (CM2) chorismate mutases.";			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=20277480; PubMed=10819329;			
RA	Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;			
RT	"Structural analysis of Arabidopsis thaliana chromosome 3. I.			
RT	Sequence features of the regions of 4,504,864 bp covered by sixty P1			
RL	and TAC clones.";			
CC	DNA Res. 7:131-135(2000).			
CC	!- CATALYTIC ACTIVITY: Chorismate = prephenate.			
CC	!- ENZYME REGULATION: Allosterically inhibited by tyrosine and			
CC	phenylalanine. Activated by tryptophan.			
CC	!- PATHWAY: BRANCH POINT OF THE BIOSYNTHETIC PATHWAY LEADING TO THE			
CC	THREE AROMATIC AMINO ACIDS, PHENYLALANINE, TYROSINE, & TRYPTOPHAN			
CC	(THE SHIKIMATE PATHWAY).			
CC	!- SIMILARITY: Contains 1 chorismate mutase domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
CC	EMBL; 226519; CAA81286.1; -.			

P40556 saccharomyc

DR EMBL; AJ242647; CAB54518.1; --  
 DR EMBL; AB026657; BAB01816.1; --  
 DR PIR; S38958; S38958.  
 DR HSP; P32178; 5CSM.  
 DR InterPro; IPR008238; Chor\_mut\_AroQ.eu.  
 DR InterPro; IPR008951; Chorismat\_mut\_II.  
 DR InterPro; IPR002701; Chorisimate\_mut.  
 DR Pfam; PF01817; Chorisimate\_mut; I.  
 DR PIRSF; PIRSF017318; Chor\_mut\_AroQ.eu; 1.  
 KW Aromatic amino acid biosynthesis; isomerase; chloroplast;  
 KW Transist peptide; Allosteric enzyme.  
 FT TRANSIT 1 60 CHLOROPLAST (POTENTIAL).  
 FT CHAIN 61 334 CHORISIMATE MUTASE.  
 FT CONFLICT 253 253 A -> R (IN REF. 1).  
 FT CONFLICT 334 AA; 37594 MW; 4D477808E349B08C CRC64;  
 SQ SEQUENCE 334 AA; 37594 MW; 4D477808E349B08C CRC64;  
 Query Match 52.1%; Score 698; DB 1; Length 334;  
 Best Local Similarity 50.4%; Pred. NO. 2.8e-47;  
 Matches 131; Conservative 46; Mismatches 75; Indels 8; Gaps 1;  
 QY 10 DSGNVYTLASVREDLAVROEDRIIYGLIERAKFPNSHTYDEKYAQIOGFCGSLVEFVVKX 69  
 Db 75 DESESLTLEGINSIREDIIIFLLERAKYCNADTYDTAFDMGFGSLVEYVWKG 134  
 QY 70 TEAIQAKAGRYKXNPENAFFPENPPSIVPSYKQFLHPGAASININKSIWQYFKELL 129  
 Db 135 TEKLHAKVGRPKSPDEHPFDDLPPEMLPQLQPKVLHPAADSININKKIWNMYFRDLV 194  
 QY 130 PLLATSGDGNVYAQTAANDLSLQISIRHYGKVAEVRDPAQDYELIRAKDEGL 189  
 Db 195 PLVKKGGDGNVYAVCAIDCLQCLSRHYGKVAEVRDPAQDYELIRAKDEGL 254  
 QY 130 MKLLFTTSVEETVRKVRKKAQVFCQVNLNSDNDNENR-----KDPSPVASSLLY 241  
 Db 255 MDMLTFTVEDAIKRVEMKTRTYQEVKVGMEKEEENSHVYKISPLVGLDYG 314  
 QY 242 NWVPLTKEVQVYLLRLD 261  
 Db 315 DWIMPLTKEVQVYLLRLD 334  
 RESULT 2  
 CHMU YEAST  
 ID CHMU YEAST STANDARD; PRT; 256 AA.  
 AC P32178;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Chorisimate mutase (EC 5.4.99.5) (CM).  
 GN ARO7 OR OSM2 OR YPR060C OR YP9499.15C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X2180;  
 RX MEDLINE=89155418; PubMed=2646272;  
 RA Schmidheini T., Sperisen P., Paravicini G., Huettner R., Braus G.H.;  
 RT "A single point mutation results in a constitutively activated and  
 feedback-resistant chorisimate mutase of Saccharomyces cerevisiae.";  
 RL J. Bacteriol. 171:1245-1253 (1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=97313271; PubMed=9169875;  
 RA Bussey H., Storms R.K., Ahmed A., Albertmann K., Allen E., Ansoorge W.,  
 RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,  
 RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,  
 RA Chung B., Churcher C.M., Coster F., Davis K., Davis R.W.,  
 RA Dietrich F.S., Dalius H., DiPaolo T., Dubois E., Duesterhoeft A.,  
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,  
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,  
 Hunicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,  
 Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,  
 Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,  
 Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,  
 Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,  
 Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,  
 Urestarazu L.A., Ushinsky S., Vierdeels F., Visers S., Voss H.,  
 Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,  
 Zhong W.W., Zollner A., Vo D.H., Hani J.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";  
 RL Nature 387:103-105 (1997).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE=95062155; PubMed=7971967;  
 RA Xue Y., Lipscomb W.N., Graf R., Schnappauf G., Braus G.;  
 RT "The crystal structure of allosteric chorisimate mutase at 2.2-A  
 resolution.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10814-10818 (1994).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=96194968; PubMed=8622937;  
 RA Straeter N., Haakansson K., Schnappauf G., Braus G., Lipscomb W.N.;  
 RT "Crystal structure of the R state of allosteric yeast chorisimate  
 mutase and comparison with the R state.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:3330-3334 (1996).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=98046093; PubMed=9384560;  
 RA Straeter N., Schnappauf G., Braus G., Lipscomb W.N.;  
 RT "Mechanisms of catalysis and allosteric regulation of yeast  
 chorisimate mutase from crystal structures.";  
 RL Structure 5:1437-1452 (1997).  
 CC -!- CATALYTIC ACTIVITY: Chorisimate = prephenate.  
 CC -!- a strong inhibitor. Allosterically regulated.  
 CC -!- PATHWAY: BRANCH POINT OF THE BIOSYNTHETIC PATHWAY LEADING TO THE  
 THREE AROMATIC AMINO ACIDS, PHENYLALANINE, TYROSINE, & TRYPTOPHAN.  
 CC -!- SUBUNIT: Homodimer.  
 CC -!- SIMILARITY: Contains 1 chorisimate mutase domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; M24517; AAB59309.1; --  
 DR EMBL; Z49219; CA889177.1; --  
 DR EMBL; Z71255; CA95004.1; --  
 DR PIR; A45921; A45921.  
 DR PDB; 1CSM; 15-SEP-95.  
 DR PDB; 2CSM; 23-DEC-96.  
 DR PDB; 3CSM; 14-JAN-98.  
 DR PDB; 4CSM; 14-JAN-98.  
 DR PDB; 5CSM; 14-JAN-98.  
 DR GeneOnline; 144325; --  
 DR SGD; S0006264; ARO7.  
 DR GO; GO:0004106; F:chorisimate mutase activity; IDA.  
 DR InterPro; IPR008238; Chor\_mut\_AroQ.eu.  
 DR InterPro; IPR008951; Chorisimat\_mut\_II.  
 DR InterPro; IPR002701; Chorisimate\_mut.  
 DR Pfam; PF01817; Chorisimate\_mut; I.  
 DR PIRSF; PIRSF017318; Chor\_mut\_AroQ.eu; 1.  
 KW Aromatic amino acid biosynthesis; isomerase; 3D-structure;  
 Allosteric enzyme.  
 FT VARIANT 226 226 T -> I (CONSTITUTIVELY ACTIVATED AND  
 FEEDBACK-RESISTANT).  
 FT TURN 3 4  
 FT HELIX 6 9  
 FT HELIX 12 34  
 FT TURN 35 35

```
FT HELIX 40 42
FT TURN 43 43
FT TURN 45 46
FT TURN 48 49
FT TURN 53 54
FT TURN 59 73
FT TURN 74 75
FT HELIX 76 78
FT TURN 80 81
FT TURN 89 89
FT TURN 106 107
FT TURN 108 110
FT HELIX 114 124
FT TURN 125 125
FT HELIX 126 129
FT TURN 137 138
FT HELIX 140 159
FT TURN 160 160
FT HELIX 161 170
FT TURN 171 171
FT HELIX 173 181
FT TURN 182 183
FT HELIX 185 191
FT TURN 192 192
FT TURN 195 211
FT TURN 212 213
FT HELIX 227 236
FT TURN 237 237
FT HELIX 238 251
FT TURN 252 252
SQ SEQUENCE 256 AA; 8C6BEBEAA3497E23 CRC64;

Query Match 37.2%; Score 498.5; DB 1; Length 256;
Best Local Similarity 40.5%; Pred. No. 7.1e-32;
Matches 102; Conservative 57; Mismatches 84; Indels 9; Gaps 4;

QY 14 VYTLASVREDLVROEDTIYGLIERAKFPSNSHTYDEKY--AQIQGFCGSLVFVVKTE 71
DB 9 VLNLQIRDELVRMEDSIIFKFIERSHFATCPVSVEANHPGLEIFNFKGSLDVALSNLE 68
QY 72 ATQAKGRYKNPEENAFFPENLPSPVSPYFKQFLHPGGAAGININKSIWKMYFKELLP 131
DB 69 IAHSTRRESPEDETFDFDKQKSLPINSYQILAPYAPEVNYNDKIKVYIEKIPL 128
QY 132 LA-TSGDD-GNYAQTAAANDLSLQISRIHYGKFAVVKFRDAPODYEPLTRAKDKEGL 189
DB 129 ISKRDGDKNNFGSVATRDIECLQSLSRRIHFGKFVAEAKFQSDIPLYTKLIKSKDVEGI 188
QY 190 MKLLTFTSVETVRKVEKKAUVFGQEVNLSDDNENRKNKFDPSVASSLYKNWVPIPTK 249
DB 189 MKNITSAAVEEKILERTKKAEGVGD-----PTNESGERRITPEVLVKIYKEIPIPTK 243
QY 250 EVQVEYLLRLD 261
DB 244 EVEVEYLLRLE 255

RESULT 3
CHMU SCHPO STANDARD; PRT; 251 AA.
AC OI3739;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable chorismate mutase (EC 5.4.99.5) (CM).
GN SPAC16E8.04C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer H.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehtach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goiffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Szpakowski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- PATHWAY: BRANCH POINT OF THE BIOSYNTHETIC PATHWAY LEADING TO THE
CC -!- CATALYTIC ACTIVITY: Chorismate = prephenate.
CC -!- SUBUNIT: Homodimer (By similarity)
CC -!- SIMILARITY: Contains 1 chorismate mutase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z98529; CAB11033.1; -.
DR PIR; T37784; T37784.
DR HSSP; P32178; 5CSM.
DR GeneDB_Spombe; SPAC16E8.04c; -.
DR InterPro; IPR008238; Chor_mut_AroQ.eu.
DR InterPro; IPR008951; Chorismat_mut_II.
DR InterPro; IPR002701; Chorismate_mut.
DR Pfam; PF01817; Chorismate_mut; 1.
DR PIRSF; PIRSF017318; Chor_mut_AroQ.eu; 1.
DR Aromatic amino acid biosynthesis; Isomerase.
KW SEQUENCE 251 AA; 29050 MW; 1AC18AE4C1B6C4B7 CRC64;

Query Match 34.5%; Score 462.5; DB 1; Length 251;
Best Local Similarity 39.5%; Pred. No. 4.5e-29;
Matches 98; Conservative 54; Mismatches 85; Indels 11; Gaps 5;

QY 17 LASVREDLVROEDTIYGLIERAKFPSNSHTY---DEKYAQIQGFCGSLVFVVKTEAI 73
DB 10 LENIRSAIQRDITIIIFNLRAQFPNEKYKSGKEGCLNLDYDGSFLNLLHEEKV 69
QY 74 QAKAGRYKNPEENAFFPENLPSPVSPYFKQFLHPGGAAS-ININKSIWKMYFKELLP 133
DB 70 VALVRYASPEEYP-FTDNLPELIPKFSKGKFLHPN--NVNVNSEILEYVINEIVPKIS 126
QY 134 TSGDD-GNYAQTAAANDLSLQISRIHYGKFAVVKFRDAPODYEPLTRAKDKEGLMKL 192
DB 127 SFGDDFDNYGTVCVDIECLQSLSRRIHYGKFAVVKFRDAPODYEPLTRAKDKEGLMKL 186
QY 193 LTFTSVETVRKVEKKAUVFGQEVNLSDDNENRKNKFDPSVASSLYKNWVPIPTKEVQ 252
DB 187 IVDAQAEERVLKELHYKALNYGRD-----AADPTKESDRINADCVASYIKDYVIPMTKKVE 242
```

QY 253 VEYLLRL 260  
DB 243 VYLLARL 250

RESULT 4  
RIR1 HAEIN STANDARD; PRT; 756 AA.  
ID R43754;  
AC P43754;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ribonucleoside-diphosphate reductase alpha chain (EC 1.17.4.1)  
DE (Ribonucleoside reductase).  
GN NRDA OR H11659.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RD / KW20 / ATCC 51907;  
RX MEDLINE=96350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Utterback T.R., Hanna M.C., Spriggs T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geohagan N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";  
RL Science 269:496-512(1995).  
CC -!- FUNCTION: Catalyzes the biosynthesis of deoxyribonucleotides from the corresponding ribonucleotides, precursors that are necessary for DNA synthesis (By similarity).  
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized thioredoxin + H(2)O = ribonucleoside diphosphate + reduced thioredoxin.  
CC -!- PATHWAY: DNA replication pathway; first step.  
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By similarity).  
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase large chain family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U32839; AAC23305.1; ALT\_INIT.  
CC HSP; P00452; 5R1R.  
CC TIGR; H11659.  
CC InterPro; IPR005144; ATP.  
CC InterPro; IPR000788; Ribonucleo\_red.  
CC InterPro; IPR008926; Ribonucleo\_red\_N.  
CC Pfam; PF03477; ATP-cone; 1.  
CC Pfam; PF00317; ribonuc\_red\_lg; 1.  
CC Pfam; PF02867; ribonuc\_red\_lgc; 1.  
CC PRINTS; PR01183; RIBORETASEM1.  
CC PROSITE; PS00089; RIBORED\_LARGE; 1.  
KW Oxidoreductase; DNA replication; Complete proteome.  
FT ACT SITE 225 225 BY SIMILARITY.  
FT ACT SITE 439 439 BY SIMILARITY.  
FT ACT SITE 462 462 BY SIMILARITY.  
FT SITE 749 749 INTERACTS WITH THIOREDOXIN/ GLUTAREDOXIN (BY SIMILARITY).

FT SITE 754 754 INTERACTS WITH THIOREDOXIN/ GLUTAREDOXIN (BY SIMILARITY).  
FT SEQUENCE 756 AA; 85695 MW; 839D02FA61E1B5D3 CRC64;  
SQ  
Query Match 7.7%; Score 103; DB 1; Length 756;  
Best Local Similarity 22.9%; Pred. No. 1.9;  
Matches 53; Conservative 36; Mismatches 104; Indels 38; Gaps 9;  
QY 49 DEKVAIQGFCGSLVEFVVK-----NTEAIQAKAGRYKNPENAFPPNLPSPVSPYS 102  
DB 231 DLSLSINATASAIIVKYSQKAGINAGAIKALGSEIRGGE--AFH-----TGCIPFYK 283  
QY 103 FKQF-----LHPGAASININKSIKWYFKELLPHATSGDDONTAQAANDLSLQ 153  
DB 284 YFTAVKSCSQGGVGGGATLY--PIWLEAENLLVKKNGVEDNRVHNDYGVQLNK 341  
QY 154 SISRIHYGKFAEVKFRDAPQDYELIRAKQEGMLKLLTTSVEETVRKVEKAVVF 213  
DB 342 LMVQRLIKGSEITLFPSPDVPLGYEAFADQDK--FEELYVKYEQDPTIRKKTVAKEIF 399  
QY 214 G---QS-----VNLNSDDNENRKFDPVASSLYKXWVPL---TKEVQ 252  
DB 400 SLLMQERASTGRYYIQNVCHCNTHSPFPDQVAPVQSNLCLEIALPTPLQ 450

RESULT 5  
SUW3 SACDO STANDARD; PRT; 737 AA.  
ID SUW3 SACDO  
AC 074727;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE ATP-dependent RNA helicase SUW3, mitochondrial precursor.  
GN SUW3.  
OS Saccharomyces douglasii (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=46617;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CK1;  
RA Golik P., Stepien P.P., Lazowska J.;  
RT "The SUW3 gene of Saccharomyces douglasii is a functional equivalent of its S. cerevisiae orthologue.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Probable ATP-dependent RNA helicase involved in a variety of mitochondrial post-transcriptional processes and in translation. It is a key control element in nuclear-mitochondrial interactions (By similarity).  
CC -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).  
CC -!- SIMILARITY: Belongs to the helicase family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AJ011586; CA09716.1;  
CC InterPro; IPR001410; DEAD.  
CC InterPro; IPR001650; Helicase\_C.  
CC Pfam; PF00271; helicase\_C; 1.  
CC SMART; SM00487; DEXDC; 1.  
CC SMART; SM00490; HELICC; 1.  
KW Mitochondrion; Transist peptide; Helicase; ATP-binding; RNA-binding.  
FT TRANSIT 1 25 MITOCHONDRION (POTENTIAL).  
FT CHAIN 26 737 ATP-DEPENDENT RNA HELICASE SUW3.  
FT NP BIND 239 246 ATP (POTENTIAL).  
SQ SEQUENCE 737 AA; 84283 MW; E24DD4C2097CA961 CRC64;  
Query Match 7.5%; Score 101; DB 1; Length 737;

Best Local Similarity 20.1%; Pred. No. 2.7; Mismatches 66; Conservative 48; Indels 121; Gaps 18;

QY 11 SGNVYTLASVREDLVRQEDTIYGLIERAKFPNSHTYDEKYAQIQGCGSLVFVVKNT 70  
DB 72 SNNVY-----LRDPLFRDN-----LDRAM---QIIYNEKLSLDA-----KQVPKQL 111  
QY 71 EAIQAKAGRYKNPENAFNPENLPSPVPSVSKQFLHPGAAS-----ININK---SIW 121  
DB 112 AWLKLKRDSTYQOLED-----PKQOTKNYVPST--BIYFSSPGNLSLILNCKNIGNSVW 165  
QY 122 K-----MYKELLPLLATSGD-DGNYAQTAAND--LS 150  
DB 166 KSLKNGQSNSTLTKPHVLVQTFDHYEOEILPMNTDDTDGAHNVDTITPAEWFS 225  
QY 151 LQOSTSRRI--HYGKFAEVRKPRDAPQDYEPILIRAKDKE-----GLMKLLFTSVSEVR 203  
DB 226 EARKIRRHIMIGP-----TNSGKYRALQKLKSVDRGYAGPLRLLA-REYIDRFQ 277  
QY 204 KRVEKXAVVFGGVNLSNDNNE-----NRKDPESVASSLY-----KXW- 243  
DB 278 SEKVRNLTGEBVIRDLDKGNPAGLISGTVEMVPINQKFDVVVLDEIQWMSDADRGA 337  
QY 244 -----VIPLTKEV 251  
DB 338 WTNALLGVVSKVHLVGEKSVLPVKSI 365

RESULT 6  
RIR1\_BUCBP STANDARD; PRT; 761 AA.  
AC Q89AS4;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Ribonucleoside-diphosphate reductase alpha chain (BC 1.17.4.1)  
DE (Ribonucleotide reductase).  
GN NRDA OR BBP168.  
OS Buchnera aphidicola (subsp. Baizongia pistaciae).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Buchnera.  
OX NCBI\_TaxID=135842;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426901; PubMed=12522265;  
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,  
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,  
RA Tamames J., Vigueria E., Iatorre A., Valencia A., Moran F., Moya A.;  
RT "Reductive genome evolution in Buchnera aphidicola."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).  
CC -I- FUNCTION: Catalyzes the biosynthesis of deoxyribonucleotides from  
CC the corresponding ribonucleotides, precursors that are necessary  
CC for DNA synthesis (By similarity).  
CC -I- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized  
CC thioresoxin.  
CC -I- PATHWAY: DNA replication pathway; first step.  
CC -I- SUBUNIT: Tetramer of two alpha and two beta chains (By  
CC similarity).  
CC -I- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase  
CC large chain family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE014016; AAC26901.1; -  
CC InterPro; IPR005144; ATP.  
CC InterPro; IPR000788; Ribonucleo\_red.

InterPro; IPR008926; Ribonucleo\_red\_N.  
DR Pfam; PF03477; ATP-cone; 1.  
DR Pfam; PF00317; ribonuc\_red\_lg; 1.  
DR Pfam; PF02867; ribonuc\_red\_lgc; 1.  
DR PRINTS; PRO1183; RIBORDTASEM1.  
DR PROSITE; PS00089; RIBORED\_LARGE; 1.  
KW Oxidoreductase; DNA replication; Complete proteome.  
FT ACT\_SITE 225 225 BY SIMILARITY.  
FT ACT\_SITE 439 439 BY SIMILARITY.  
FT ACT\_SITE 462 462 BY SIMILARITY.  
FT ACT\_SITE 754 754 INTERACTS WITH THIOREDOXIN/ GLUTAREDOXIN  
FT SITE 759 759 (BY SIMILARITY)  
FT SITE 759 759 INTERACTS WITH THIOREDOXIN/ GLUTAREDOXIN  
FT SITE 759 759 (BY SIMILARITY).  
SQ SEQUENCE 761 AA; 87124 MW; 7FB14E136283BDCF CRC64;

Query Match 7.1%; Score 94.5; DB 1; Length 761;  
Best Local Similarity 20.8%; Pred. No. 9;  
Matches 46; Conservative 42; Mismatches 104; Indels 29; Gaps 8;

QY 49 DEKYAQIQGCGSLVFVVKNTAIOAKAGRYK---NPEENAFNPENLPSPVPSVSKQ 105  
DB 231 EDNLNSINATTSALVKVYSQRA-GIGINAGRIRALGPSIRNG---DTLHTGCIFFYKHQ 286  
QY 106 F-----LHPCAASININKSIWYKELLPLLATSGDGNYAQTAANDLSLQSI 156  
DB 287 SAVKSCSGGVGGGAATIFY--PIWHFEIESLLVLKNRGIENRVRHMDYSVQLNKLNY 344  
QY 157 RRIHYGKFAEVRKPRDAPQDYEPILIRAKDKEGLMKLLTFTSVETVKRVEKXAVWFG-- 214  
DB 345 QRLILGEHILTFSPSDVPLNDYGFNNQEK--PERLYIKYENDKTIKKVKASYLFSLM 402  
QY 215 -QE-----VNLNSDDNENRKPDPSPVASSLYKNWVPLT 248  
DB 403 MQERTSTGRIYQNVHCHNSHSAFNPKIPIROSNLCLEIT 443

RESULT 7  
YFLL\_YEAST STANDARD; PRT; 295 AA.  
AC P43613;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical 34.2 kDa protein in SAPI55-YMR31 intergenic region  
DE precursor.  
GN YFR041C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=S288c / AB972;  
RX MEDLINE=95400292; PubMed=7670463;  
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,  
RA Yamazaki M., Tashiro H., Eki T.;  
RT "Analysis of the nucleotide sequence of chromosome VI from  
RT Saccharomyces cerevisiae."  
RL Nat. Genet. 10:261-268(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=S288c / AB972;  
RX MEDLINE=96287652; PubMed=8686379;  
RA Eki T., Naitou M., Hagiwara H., Ozawa M., Sasanuma S.-I.,  
RA Sasanuma M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.;  
RT "Analysis of a 36.2 kb DNA sequence including the right telomere of  
RT chromosome VI from Saccharomyces cerevisiae."  
RL Yeast 12:149-167(1996).  
CC -I- SIMILARITY: Contains 1 J domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration





RN SEQUENCE FROM N.A.  
RA Wedler H., Scharfe M., Wedler E., Wambutt R.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN  
RP SEQUENCE OF 1-19 FROM N.A.  
RA Hernandez K., Weber N., Wipfli P., Schmidheini T.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN  
RP SUBCELLULAR LOCATION, AND FUNCTION.  
EX MEDLINE=22016919; PubMed=12022229;  
RA Oeffinger M., Lueng A., Lamond A., Tollervey D.;  
RT "Yeast Pescadillo is required for multiple activities during 60S  
ribosomal subunit synthesis.";  
RL RNA 8:626-636(2002).  
CC  
CC -!- FUNCTION: May function both in ribosome synthesis and in cell-  
cycle regulation.  
CC  
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.  
CC  
CC -!- SIMILARITY: Contains 1 BRCT domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; 272888; CAA97106.1; -.  
DR PIR; S64410; S64410.  
DR Geronline; 141415; -.  
DR SGD; S0003335; RRP43.  
DR GO; GO:0005730; C:nucleolus; IDA.  
DR GO; GO:0042273; P:ribosomal large subunit biogenesis; IDA.  
DR InterPro; IPR001357; BRCT.  
DR Pfam; PF00533; BRCT; 1.  
DR SMART; SM00292; BRCT; 1.  
DR PROSITE; PS01072; BRCT; 1.  
KW Ribosome biogenesis; Nuclear protein; Coiled coil.  
FT DOMAIN 291 346  
FT DOMAIN 355 449  
FT DOMAIN 526 604  
FT DOMAIN 560 580  
FT DOMAIN 479 496  
FT DOMAIN 501 515  
FT DOMAIN 501 515  
SQ SEQUENCE 605 AA; 69877 MW; 131001C956787BE5 CRC64;  
  
Query Match  
Best Local Similarity 6.9%; Score 92; DB 1; Length 605;  
Matches 62; Conservative 50; Mismatches 116; Indels 74; Gaps 13;  
  
QY 2 AKAAEQSPGNGVYTLASVRE-----DLVRQEDTIYGLIERAKPPNSHTYDEKYAQ 54  
DB 102 AKRIEENRDS---YTLDHIIKERYPSPFPDAIRDIDALNMLFLFSLPSTNOVSSKIIND 158  
QY 55 IQGFCGSLVFVVKV-----TEAIQAKAGRYKPNPEAFPEPLPSI-- 97  
DB 159 AQKICNQLAVAKERLVKVFVSIKVVYQANIKGEVRLVLPFK---FPENISDVDF 215  
QY 98 VPSYFVKQFLHPGAASININKSIWMYFKBLPLLATSGDGNAYAGTANDLSLQISIR 157  
DB 216 RIMLTFLFELY-----STLLHFVLYKLY-----TDSGLIYPPKLDLKKDKIIS- 257  
QY 158 RIHYGKFAEVKFRDAPQDVEPL-----IRAKKSGELMKLLTF-----TSVEETR 203  
DB 258 --GLSSVILSRQDSLLKLPDTEIEDVKVSLDASTLSALNADENATDETEKEE 315  
QY 204 KRVEK-KAVFEGQVNLNDDNNEN-----KFDPSVASSLYKNWVILTKVEQVE 254  
DB 316 KKQKEQKEQNEETELDTEDNNKNGDILIQSKYDSPA-SLPSAFVYVVSREVPID 374  
QY 255 YL 256

DB 375 IL 376  
RESULT 10  
LON BRUME  
ID - LON BRUME STANDARD; PRT; 812 AA.  
AC Q8VHC6;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE ATP-dependent protease La (EC 3.4.21.53).  
GN LON OR BMEI0876.  
OS Brucella melitensis.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Brucellaceae; Brucella.  
OX NCBI\_TaxID=23459;  
RN  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=16W / ATCC 23456 / Biotype 1;  
RX MEDLINE=20020109; PubMed=11756688;  
RA DelVecchio V.G., Kapural V., Redkar R.J., Patra G., Mujar C., Los T.,  
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,  
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,  
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
RA Haselkorn R., Kyrides N., Overbeek R.;  
RT "The genome sequence of the facultative intracellular pathogen  
Brucella melitensis.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).  
CC  
CC -!- FUNCTION: Degrades short-lived regulatory and abnormal proteins in  
presence of ATP. Hydrolyzes two ATPs for each peptide bond cleaved  
in the protein substrate (By similarity).  
CC  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of large proteins such as globin,  
casein and denaturated serum albumin, in presence of ATP.  
CC  
CC -!- SUBUNIT: Homotetramer (By similarity).  
CC  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC  
CC -!- SIMILARITY: Belongs to peptidase family S16.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE009528; AAL52057.1; ALT\_INIT.  
DR PIR; AF3361; AF3361.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR003593; AAA\_ATPase\_cent.  
DR InterPro; IPR001270; Chaprin\_c1pA/S.  
DR InterPro; IPR001984; Peptidase\_S16.  
DR InterPro; IPR008268; Peptid\_S16\_AS.  
DR InterPro; IPR008269; Peptid\_S16\_C.  
DR InterPro; IPR004815; Pept\_S16\_Lon.  
DR InterPro; IPR003111; Pept\_S16\_N.  
DR Pfam; PF00004; AAA; 1.  
DR Pfam; PF02190; LON; 1.  
DR Pfam; PF05362; Lon\_C; 1.  
DR PRINTS; PR00300; CLPPROTEASEA.  
DR PRINTS; PR00830; ENDOLAPTAASE.  
DR SMART; SM00382; AAA; 1.  
DR SMART; SM00464; LON; 1.  
DR TIGRFAMs; TIGR00763; Lon; 1.  
DR PROSITE; PS01046; LON\_SER; 1.  
KW Hydrolase; Serine protease; ATP-binding; Complete proteome.  
FT NP\_BIND 369 374 ATP (POTENTIAL).  
FT ACT\_SITE 689 689 BY SIMILARITY.  
SQ SEQUENCE 812 AA; 93895 MW; 401E15F1FDD24E98 CRC64;  
  
Query Match  
Best Local Similarity 6.9%; Score 92; DB 1; Length 812;  
Matches 63; Conservative 44; Mismatches 85; Indels 94; Gaps 14;

QY 4 AAEQSPDSGNVY---TLASVREDLVROEDTIYGLIE---RAKFP--SNSHTYDEKYAQI 55  
 DB 69 AADDDPAPDAIYEIGTIANVLQ--LLKLPDGTVKVLVEGTARAKISKFTDREDYHEAYA-- 125  
 QY 56 QGFCGSLVFEVVKNTAIOAKAGRYKNPEENAFFPENLPPSVPSYKQFLHPGAAAIN 115  
 DB 126 -----AALQ-----EPEDAVEIEALARSVVS--DFENY-----VK 154  
 QY 116 INKSIWMTFKKELLPLLATSGDGNVAQTAANDLSLOSISRRIHYGKFAEVKFRDAPQ 175  
 DB 155 LNKKI-----SPEVGTASQIDDDYSLADTVASHLAI-----KIPE 190  
 QY 176 DYEPIIRAKDKKGLMKLLTF-----TSVETVKRVEKKAIVFGQVWNLN----- 220  
 DB 191 KOEMLSVLSVRERLEKALSFMEAEISVLQVKRISRVKQMKETQREYYLNEQWKAIOK 250  
 QY 221 -----SDDNDNENRKFPDPSVASSLYKNWVPLTKVQVEYL--LRRL 260  
 DB 251 ELGDSGDRGEVAEIEERTIKT-----KLSKEAREKALAEKKL 289

RESULT 11  
 LON\_BRUSU STANDARD; PRT; 812 AA.  
 AC Q8G017;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DE ATP-dependent protease La (EC 3.4.21.53).  
 GN LON OR BR1106.  
 OS Brucella suis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29461;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1330 / Biovar 1;  
 RX MEDLINE=2247741; PubMed=12271122;  
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,  
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,  
 RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,  
 RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J., Van Aken S.E.,  
 RA Riedmuller S., Teddlin H., Gill S.R., White O., Salzberg S.L.,  
 RA Hoover D.B., Linder L.E., Halling S.M., Boyle S.M., Fraser C.M.,  
 RT "The Brucella suis genome reveals fundamental similarities between  
 animal and plant pathogens and symbionts."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).  
 CC -!- FUNCTION: Degrades short-lived regulatory and abnormal proteins in  
 presence of ATP. Hydrolyzes two ATPs for each peptide bond cleaved  
 in the protein substrate (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of large proteins such as globin,  
 casein and denaturated serum albumin, in presence of ATP.  
 CC -!- SUBUNIT: Homotetramer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to peptidase family S16.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC ENBL; AE014412; AA30026.1; -  
 DR TIGR; BR1106; -  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR003593; AAA\_ATPase.centr.  
 DR InterPro; IPR001270; Chaprinin c1pA/B.  
 DR InterPro; IPR001984; Peptidase S16.  
 DR InterPro; IPR008268; Peptid S16\_AS.  
 DR InterPro; IPR008268; Peptid S16\_C.  
 DR InterPro; IPR004815; Pept\_S16\_Lon.

InterPro; IPR003111; Pept\_s16\_N.  
 DR Pfam; PF00004; AAA; 1.  
 DR Pfam; PF02190; LON; 1.  
 DR Pfam; PF05362; Lon\_C; 1.  
 DR PRINTS; PRO0300; CLP\_PROTEASEA.  
 DR PRINTS; PRO0830; ENDOLAPTASE.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00464; LON; 1.  
 DR TIGRFAMs; TIGR00763; Lon; 1.  
 DR PROSITE; PS01046; LON\_SER; 1.  
 KW Hydrolyase; Serine protease; ATP-binding; Complete proteome.  
 FT NP\_BIND 367 374 ATP (POTENTIAL).  
 FT ACT\_SITE 689 689 BY SIMILARITY.  
 SQ SEQUENCE 812 AA; 89794 MW; 0FOAC318C9227FBF CRC64;  
 Query Match 6.9%; Score 92; DB 1; Length 812;  
 Best Local Similarity 22.0%; Pred. No. 15;  
 Matches 63; Conservative 44; Mismatches 85; Indels 94; Gaps 14;  
 QY 4 AAEQSPDSGNVY---TLASVREDLVROEDTIYGLIE---RAKFP--SNSHTYDEKYAQI 55  
 DB 69 AADDDPAPDAIYEIGTIANVLQ--LLKLPDGTVKVLVEGTARAKISKFTDREDYHEAYA-- 125  
 QY 56 QGFCGSLVFEVVKNTAIOAKAGRYKNPEENAFFPENLPPSVPSYKQFLHPGAAAIN 115  
 DB 126 -----AALQ-----EPEDAVEIEALARSVVS--DFENY-----VK 154  
 QY 116 INKSIWMTFKKELLPLLATSGDGNVAQTAANDLSLOSISRRIHYGKFAEVKFRDAPQ 175  
 DB 155 LNKKI-----SPEVGTASQIDDDYSLADTVASHLAI-----KIPE 190  
 QY 176 DYEPIIRAKDKKGLMKLLTF-----TSVETVKRVEKKAIVFGQVWNLN----- 220  
 DB 191 KOEMLSVLSVRERLEKALSFMEAEISVLQVKRISRVKQMKETQREYYLNEQWKAIOK 250  
 QY 221 -----SDDNDNENRKFPDPSVASSLYKNWVPLTKVQVEYL--LRRL 260  
 DB 251 ELGDSGDRGEVAEIEERTIKT-----KLSKEAREKALAEKKL 289

RESULT 12  
 ATPA\_BUCAP STANDARD; PRT; 510 AA.  
 AC O51874;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ATP synthase alpha chain (EC 3.6.3.14).  
 GN ATPA OR BUSG006.  
 OS Buchnera aphidicola (subsp. Schizaphis graminum).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Buchnera.  
 OX NCBI\_TaxID=98794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97361981; PubMed=9216881;  
 RA Clark M.A., Baumann P.;  
 RT "The (P1FO) ATP synthase of Buchnera aphidicola (endosymbiont of  
 aphids): genetic analysis of the putative ATP operon."  
 RL Curr. Microbiol. 35:84-89(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98184963; PubMed=9516544;  
 RA Clark M.A., Baumann L., Baumann P.;  
 RT "Sequence analysis of a 34.7-kb DNA segment from the genome of  
 Buchnera aphidicola (endosymbiont of aphids) containing groEL, dnaA,  
 RT the atp operon, gldA, and rho."  
 RL Curr. Microbiol. 36:158-163(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22084549; PubMed=12089438;  
 RA Tamás I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,  
 RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;

## RESULT 14

ALP6 SCHPO STANDARD; PRT; 832 AA.  
 ID ALP6 SCHPO STANDARD; PRT; 832 AA.  
 AC Q9USQ2; Q94366; Q9P954;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Spindle pole body component alp6 (Altered polarity protein 6).  
 GN ALP6 OR SPBC428.20C OR SPBC902.01C.  
 OS Schizosaccharomyces pombe (Fission Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RP MEDLINE=20532503; PubMed=11080156;  
 RA Vardy L., Toda T.;  
 RT "The fission yeast gamma-tubulin complex is required in G(1) phase and  
 RT is a component of the spindle assembly checkpoint.";  
 RL EMBO J. 19:6098-6111(2000).  
 RN [2]  
 SEQUENCE FROM N.A.  
 RP STRAIN=972;  
 RC MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Collins M., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quayle M.A., Rabinowitsch E.,  
 RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J.K., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Welter J., Vanstreels E., Rieger M., Schaefer M., Moestl D., Hilbert H.,  
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huxst S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: Component of the gamma tubule complex that is required  
 CC for the regulation of both interphase microtubules and mitotic  
 CC bipolar spindles.  
 CC -!- SUBCELLULAR LOCATION: Spindle pole body and the microtubule  
 CC organizing center (MTOC).  
 CC -!- SIMILARITY: Belongs to the GCP family.  
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to a  
 CC frameshift in position 1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB040811; BA294097.1; ALT\_FRAME.  
 CC EMBL; AL033306; CA562095.1; -  
 CC EMBL; AL034382; CA22295.1; -  
 CC GeneDB SPombe; SPBC428.20C; -  
 CC InterPro; IPR007259; SPC97\_Spc98.  
 CC Pfam; PF04130; SPC97\_Spc98; 1.  
 CC -----  
 CC KW Chapterone; ATP-binding; Complete proteome.  
 CC NP\_BIND 56 63 ATP (POTENTIAL).  
 CC FT

Microtubule; Mitosis.  
 FT CONFLICT 665 675 MISSING (IN REF. 2).  
 SQ SEQUENCE 832 AA; 95996 MW; C92771C3DBF5COLA CRC64;  
 Query Match 6.8%; Score 91.5; DB 1; Length 832;  
 Best Local Similarity 25.4%; Pred. No. 17;  
 Matches 45; Conservative 28; Mismatches 67; Indels 37; Gaps 11;  
 QY 108 HPGNASININKSWK-MYF---KELLPLLATL-----GDDGVYAQTAAANDLSLQSI 155  
 Db 408 HQGGQDV-----VWKGKYLKELIPSLSEELVDKIFLIGKSLNFARYCGGDFDWAQE- 461  
 QY 156 SRRHYGKFAEVKFRDAPQDYPRAKDKELMKLLTFTSVEE---TVRKRVKKAV 211  
 Db 462 ----HYQKLVKLSYRD-FHSLETVDVKAYTESINHLVYL--MEEVFHLTDHLKAKKYL 514  
 QY 212 VFQGE--VNINSDNDNE-----NRKDPDSVASSLYKNWVILPTKEVQVEYLLRLD 261  
 Db 515 LLGGQDFVDLLMESLGNLSDQPANTLFRHNLTAFL-ESAIRSSNASEYEPEYVLKRLD 570

RESULT 15  
 HSLU CAMJE STANDARD; PRT; 439 AA.  
 ID HSLU CAMJE STANDARD; PRT; 439 AA.  
 AC Q9PHLO;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE ATP-dependent hsl protease ATP-binding subunit hslu.  
 GN HSLU OR CJ0662C.  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Campylobacteraceae; Campylobacter.  
 OX NCBI\_TaxID=197;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RP STRAIN=NCCT 11168;  
 RC MEDLINE=20150912; PubMed=10689204;  
 RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,  
 RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
 RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
 RA Whitehead S., Barrell B.G.;  
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
 RT reveals hypervariable sequences.";  
 RL Nature 403:665-668(2000).  
 CC -!- FUNCTION: Chaperone subunit of a proteasome-like degradation  
 CC complex (By similarity).  
 CC -!- SUBUNIT: A double ring-shaped homohexameric of hslu is capped on  
 CC each side by a ring-shaped hslu homohexamer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the clpX chaperone family. Hslu subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AL139075; CAB75294.1; -  
 CC PIR; H81414; H81414.  
 CC HSP; P32168; 1D02.  
 CC HAMAP; MF 00249; - 1.  
 CC InterPro; IPR003593; AAA\_ATPase.  
 CC InterPro; IPR003959; AAA\_ATPase\_Centr.  
 CC InterPro; IPR004491; Hsp\_Hslu.  
 CC Pfam; PF00004; AAA; 1.  
 CC SMART; SM00382; AAA; 1.  
 CC TIGRFAMs; TIGR00390; hslu; 1.  
 CC Chapterone; ATP-binding; Complete proteome.  
 CC NP\_BIND 56 63 ATP (POTENTIAL).  
 CC FT

SQ SEQUENCE 439 AA; 49779 MW; 36FE3914DFBE48DA CRC64;  
Query Match 6.8%; Score 91; DB 1; Length 439;  
Best Local Similarity 25.2%; Pred. No. 8.5;  
Matches 53; Conservative 31; Mismatches 76; Indels 50; Gaps 11;  
QY 51 KYAQIQGFCGSLVEFVVKNTAIAQAG-----RYKNPEENAFFPENLPPSIVPSYS 102  
DB 84 KYTEV-GFVGRDVESMRDL---ANAALNLVKNQREKKNKIDEFIEN-----KI 130  
QY 103 FKQFLHFGAASININKSIWNYFKELLPLLAT---SGD-----DGNVAQTAAN-- 147  
DB 131 LEKLLPPLPKGISDEK---QSEYKNSLEKQRTKLRNGDLDESTIEIISQNWFDINPILP 187  
QY 148 -DLSLLQISRRRIHYG--KFVAEVKFRDAPQDYELIRAKXEGMLKLLTFTSVETVRK 204  
DB 188 PENGAMQDIVKIVGVGSKVKVKKEMKIDAK-----NALKNEAGEKILDQESIKSEALK 240  
QY 205 RVEKKAVFQGVNLSNDNDNENKFDPS 234  
DB 241 RAENEGIIIFIDEIDKIAVSSGNSNRQ-DPS 269

Search completed: September 25, 2004, 02:53:01  
Job time : 20 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 25, 2004, 02:55:17 ; Search time 4193 Seconds

(without alignments)  
2697.957 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKAAEQSPDSGNVYTLASV.....NWVPLTKVQVEYLLRLD 261

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+pn.model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US10624061/runat\_24092004\_092406\_16911/app query.fasta\_1.455  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptO -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10624061@cgn 1.1 3731 @runat\_24092004\_092406\_16911 -NCPU=6 -ICPU=3  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.scs:\*  
28: em.un:\*

29: em.vi:\*  
30: em.htg.hum:\*  
31: em.htg.inv:\*  
32: em.htg.other:\*  
33: em.htg.mus:\*  
34: em.htg.pln:\*  
35: em.htg.rod:\*  
36: em.htg.mam:\*  
37: em.htg.vrt:\*  
38: em.sy:\*  
39: em.htgo.hum:\*  
40: em.htgo.mus:\*  
41: em.htgo.other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1339	100.0	1020	6	AR404688 Sequence
2	818	61.1	798	6	AX506277 Sequence
3	818	61.1	823	8	AY133840 Arabidops
4	818	61.1	993	6	AX343935 Sequence
5	818	61.1	1006	6	AR236634 Sequence
C	818	61.1	1006	6	AR236640 Sequence
7	818	61.1	1006	6	AX343933 Sequence
8	818	61.1	1006	6	AX464575 Sequence
C	818	61.1	1006	6	AX464581 Sequence
10	818	61.1	1006	6	L47355 Arabidopsis
11	818	61.1	1055	8	AY052238 Arabidops
12	767.5	57.3	525	6	AR404683 Sequence
13	698	52.1	1005	6	AX506042 Sequence
14	696	52.0	1207	6	AR236633 Sequence
C	696	52.0	1207	6	AR236639 Sequence
16	696	52.0	1207	6	AX464574 Sequence
C	696	52.0	1207	6	AX464580 Sequence
18	696	52.0	1207	8	ATCHMUT
19	691	51.6	1000	8	L47356 Lycopersico
20	686	51.2	1348	8	AK105512 Oryza sat
21	685.5	51.2	1231	6	AR404687 Sequence
22	685.5	51.2	1250	8	AK068983 Oryza sat
23	678.5	50.7	1435	8	AK101220 Oryza sat
24	677.5	50.6	1274	8	AY089156 Arabidops
25	670.5	50.1	1223	6	AR404686 Sequence
26	669.5	50.0	951	8	BT005306 Arabidops
27	669.5	50.0	1143	8	AK117860 Arabidops
28	654.5	48.9	1217	6	AR236635 Sequence
C	654.5	48.9	1217	6	AR236641 Sequence
30	654.5	48.9	1217	6	AX464576 Sequence
C	654.5	48.9	1217	6	AX464582 Sequence
32	654.5	48.9	1217	8	AF131219 Arabidops
33	646	48.2	1043	8	AK069725 Oryza sat
34	615	45.9	81062	2	AC128660 Medicago
35	602	45.0	780	6	AR404689 Sequence
36	602	45.0	780	6	BT009499 Triticum
37	587	43.8	5176	6	AR236637 Sequence
38	587	43.8	5176	6	AX464578 Sequence
39	587	43.8	5177	8	AT242648 Arabidops
C	573.5	42.8	98473	8	ATT30N20 Arabidops
41	549.5	41.0	5099	6	AR236636 Sequence
42	549.5	41.0	5099	6	AX464577 Sequence
43	549.5	41.0	5099	8	ATH242647 Arabidops
C	549.5	41.0	77401	8	AB026657 Arabidops
C	540	40.3	93802	8	AC018364 Arabidops

# ALIGNMENTS

RESULT 1

AR404688  
 LOCUS AR404688 1020 bp DNA linear PAT 18-DEC-2003  
 DEFINITION Sequence 15 from patent US 6627798.  
 ACCESSION AR404688  
 VERSION AR404688.1 GI:40153356  
 KEYWORDS Unknown.  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 1020)  
 AUTHORS Falco, S.C., Fancdu, O.O. and Lee, J.-M.  
 TITLE Aromatic amino acid biosynthetic enzymes  
 JOURNAL Patent: US 6627798-A 15 30-SEP-2003;  
 FEATURES  
 source  
 1. 1020  
 /locus\_tag="unknown"  
 /mol\_type="genomic DNA"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,57e-121 Length: 1020  
 Score: 1339.00 Matches: 261  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-624-061-16 (1-261) x AR404688 (1-1020)  
 QY 1 MetAlaLysAlaAlaGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
 Db 23 ATGGCCAAAGCAGCAGAAACAAGTCTGATTCGGAAATGTGTACACGCTAGCTTCTGTG 82  
 QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluAlaLys 40  
 Db 83 AGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTATGTCCTCATTTGAGAGGCCAAG 142  
 QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
 Db 143 TTCCTTAGCAATCTCAACACCTATGATGAAAGTATGCTCAATCCAGGGTTTTGTGGC 202  
 QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
 Db 203 TCATTGGTGGATTTGTTTAAAGATACAGAGGCCATTCAGCTAAGCTAAGCTGGAAGATAC 262  
 QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100  
 Db 263 AAAAACCCCTGAAGAAACCGCTTCTTCCAGAAAATTTACCACCATCAATTGTGCCATCT 322  
 QY 101 TyrSerPheLysGlnPheLeuHisProGlyValAlaSerIleAsnIleAsnLysSerIle 120  
 Db 323 TACTCCTTCAACACAGTTTTTCATCCTCGTCTGCTCTCAATTAACAATAAACAAGTCCATC 382  
 QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsn 140  
 Db 383 TGGAAAATGTATTTCAAGAGTTACTTCCATTGCTTCTACTTCTCGGTGATGATGCCAAC 442  
 QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgAlaIleHis 160  
 Db 443 TATGGGCAACTGCAGCTTAATGACCTTTTCATTATTTGACGTCCTCTCTAGAAGATTCCAC 502  
 QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
 Db 503 TATGGAAAATTTGTAGCTAGGTGAATTCAGGGATCTCTCAAGACTACGAGCTTTTA 562  
 QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200  
 Db 563 ATTCGAGCTAAGGATAAAGAGGATGATGAATTTGTGACATTTAACAAGCGTTGAAGAG 622  
 QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAlaSerLeuAsn 220  
 Db 623 ACGGTGAGGAAGAGAGTTGAAGAAGAGGCTGTGCTGTTTGGCAGGAGAGTGAATCTTAAC 682

QY 221 SerAspAspAsnAspAsnGluAsnArgLysPheAspProSerProSerValAlaSerSerLeuTyr 240  
 Db 683 AGTGATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC 742  
 QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGlnValGlnValGlnValGlnVal 260  
 Db 743 AAAAATTGGGTGATACCTCTCACCAAGAGGTTTCAGGTTGAGTACCTCTTCGCGCGCTCTA 802  
 QY 261 Asp 261  
 Db 803 GAC 805  
 RESULT 2  
 AX506277  
 LOCUS AX506277 798 bp DNA linear PAT 27-SEP-2002  
 DEFINITION Sequence 972 from Patent WO0216655.  
 ACCESSION AX506277  
 VERSION AX506277.1 GI:23387514  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE 1  
 AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.  
 TITLE Stress-regulated genes of plants, transgenic plants containing  
 same, and methods of use  
 JOURNAL Patent: WO 0216655-A 972 28-FEB-2002;  
 The Scripps Research Institute (US); Syngenta Participations AG  
 (CH)  
 FEATURES  
 Location/Qualifiers  
 source  
 1. 798  
 /organism="Arabidopsis thaliana"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3702"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.29e-70 Length: 798  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservatives: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 6 Gaps: 2  
 US-10-624-061-16 (1-261) x AX506277 (1-798)

QY 1 MetAlaLysAlaAlaGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 Db 1 ATGGCAAGAGTCTTCCGAATCGGATTCGGGTTCGTGTTCCCAATGACTCAGCTTGAC 60  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluArg 38  
 Db 61 TTAATCAGAGATTCGTTGATTAGGCAAGACACCATCGTCTTCAGCTTCATCGAGAGA 120  
 QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 Db 121 GCTAAGTTTCCATCAATTCCTGCTTCGAGGAATTCGTTGCTAGATTCGGAAGT 180  
 QY 59 CysGlySerLeuValGluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 Db 181 TTCTCTTCTCCTAGTCTTTCGTCAGAGACAGAAATCATCCAGCTTAAGGTAGGA 240  
 QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleVal 98  
 Db 241 AGATATGAATACCCGGAAGAGAAATCTTCTTCCTTCAGAACATTCCTCACTCGGTTTT 300  
 QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
 Db 301 CCTACGACACAATAATCATCGGCTTTCACCCCAAGGCTCTATCTGTAAACATTAACAAA 360  
 QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138







SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1006)  
AUTHORS Boyes,D.C., Davis,K.R., Woessner,J.P., Goriach,J., Hamilton,C.M.,  
Hoffman,N.E., Kloti,A.S., Zayed,A. and Ascenzi,R.A.  
TITLE Methods and compositions for the modulation of chorismate synthase  
and chorismate mutase expression or activity in plants  
JOURNAL Patent: US 6465217-A 3 15-OCT-2002;  
FEATURES Location/Qualifiers  
source 1..1006  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 1,73e-70 Length: 1006  
Score: 818.00 Matches: 162  
Percent Similarity: 72.83% Conservative: 31  
Best Local Similarity: 61.13% Mismatches: 68  
Query Match: 61.09% Indels: 4  
DB: 6 Gaps: 2

US-10-624-061-16 (1-261) x AR236634 (1-1006)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
DB 64 ATGGCAAGAGCTTCGAATCGGATTCGGTTCTGTTTCCAAATGATCTGAGTCTTGAC 123  
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArg 38  
DB 124 TTAATCAGAGATCGTTGATTAGGCAAGACACCATCGTCTTCAGTTGATCGAGAGA 183  
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
DB 184 GCTAAGTTCCCACTCAATCTCTGCTTTCGAGGAATCTCGTTGTCTAGATCTCGAAGT 243  
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
DB 244 TTCTCTTCTCTACTGAGTTTTCGTCAGAGACAGAAATCATCCAACTAGGTAGGA 303  
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProSerIleVal 98  
DB 304 AGATATGAATACCCGGAAGAGATCCTTCTCTTTCGAGAACATTCCTCACTCGGTTTTT 363  
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLys 118  
DB 364 CCTACGCACAAATATCCATCGGCTTTTGCACCCCTAAGGCTCTATCTGTTAACATTAA 423  
QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138  
DB 424 CAATCTGGGATATTACTTTAAGAAATTCCTTCTTGTTCGAGAACCTGGCGATGAT 483  
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArg 158  
DB 484 GGCAACTATCCATCACTGCTGATGATCTCGCTGTTTACAACTCTCTTCGAGAAGG 543  
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
DB 544 ATTCACTACCGTAAATTTAGCTGAGTCAATTCAGAGATGCTCCCAAGATTACGAG 603  
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198  
DB 604 CTTGCGATTCCGCTCAGGATAGAGAGGCTTTGATGAAGCTGTGACGTTTGAGAAAGTA 663  
QY 199 GluGluThrValArgLysArgValGluLysValAlaValPheGlyGlnGluValAsn 218  
DB 664 GAGAAATGTTTAAAGAGAGGTGCGAAGAAAGCAGAAACGTTTGGACAAAGATGAAA 723  
QY 219 LeuAsnSerAspAsnAsnGluAsnArgLysPhe-----AspProSerValAla 236  
DB 724 TTCAACTCTGCTATCGCGATGAGAGTAAAGAGAGATATAAAGTGGATTCATTGCTGCC 783

QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
DB 784 TCTCGCATCTACGGGAATGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTT 843  
QY 257 LeuArgArgLeuAsp 261  
DB 844 CTACGTCGTCCTCGAT 858  
RESULT 6  
AR236640/c  
LOCUS AR236640 1006 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 9 from patent US 6465217.  
ACCESSION AR236640  
VERSION AR236640.1 GI:27280726  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1006)  
AUTHORS Boyes,D.C., Davis,K.R., Woessner,J.P., Goriach,J., Hamilton,C.M.,  
Hoffman,N.E., Kloti,A.S., Zayed,A. and Ascenzi,R.A.  
TITLE Methods and compositions for the modulation of chorismate synthase  
and chorismate mutase expression or activity in plants  
JOURNAL Patent: US 6465217-A 9 15-OCT-2002;  
FEATURES Location/Qualifiers  
source 1..1006  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 1,73e-70 Length: 1006  
Score: 818.00 Matches: 162  
Percent Similarity: 72.83% Conservative: 31  
Best Local Similarity: 61.13% Mismatches: 68  
Query Match: 61.09% Indels: 4  
DB: 6 Gaps: 2

US-10-624-061-16 (1-261) x AR236640 (1-1006)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
DB 943 ATGGCAAGAGCTTCGAATCGGATTCGGTTCTGTTTCCAAATGATCTGAGTCTTGAC 884  
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArg 38  
DB 883 TTAATCAGAGATCGTTGATTAGGCAAGACACCATCGTCTTCAGTTGATCGAGAGA 824  
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
DB 823 GCTAAGTTCCCACTCAATTCCTCTGCTTTCGAGGAATCTCGTTGTCTAGATTCCTGGAAGT 764  
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
DB 763 TTCTCTTCTCTACTGAGTTTTCGTCAGAGACAGAAATCATCCAACTAGGTAGGA 704  
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProSerIleVal 98  
DB 703 AGATATGAATACCCGGAAGAGATCCTTCTTCTTCGAGAACATTCCTCACTCGGTTTTT 644  
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLys 118  
DB 643 CCTACGCACAAATATCCATCGGCTTTGACCCCTAAGGCTCTATCTGTTAACATTAA 584  
QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138  
DB 583 CAATCTGGGATATTACTTTAAGAAATTCCTTCTTGTTCGAGAACCTGGCGATGAT 524  
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArg 158  
DB 523 GGCAACTATCCATCACTGCTGATGATCTCGCTGTTTACAACTCTCTTCGAGAAGG 464  
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178

```

Db      463 ATTCACTACGGTAAATTTGTAGCTGAGGTCAAAATTCAGAGATGCTCCCAAGATTACGAG 404
QY      179 ProLeuileArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198
Db      403 CTGCGGATTCGGCTCAGATAGAGAGGCTTTGATGAGAGCTGTGACGCTTTGAGAAAGTA 344
QY      199 GluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsn 218
Db      343 GAAGAAATGTTTAAAGAGAGAGTGCAGAGAAAGAGAAACGTTTGGCAAGAAAGTAAA 284
QY      219 LeuAsnSerAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236
Db      283 TTCAACTCTGGCTATGGCGATCAGAGTAAGAGAGATATAAGTGCATCCATGCTTGCC 224
QY      237 SerSerLeuTyLysAsnTrpValPheProLeuThrLysGluValGlnValGluTyLys 256
Db      223 TCTCGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTT 164
QY      257 LeuArgArgLeuAsp 261
Db      163 CTACGTCGTCGAT 149

RESULT 7
AX343933      1006 bp      DNA      linear      PAT 01-FEB-2002
LOCUS
DEFINITION
Sequence 3 from Patent WO0200901.
ACCESSION
AX343933
VERSION
AX343933.1 GI:18491972
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
Badur, R., Geiger, M., Kunze, I. and Sommer, S.
Changing the fine chemical content in organisms by genetically
modifying the shikimate pathway
Patent: WO 0200901-A 3 03-JAN-2002;
Sungene GmbH & Co. KGaA (DE)
FEATURES
source
1..1006
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
64..861
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD22398.1"
/db_xref="GI:18491973"
/db_xref="REMBL:CAD22398"
/translation="MARVFESDGGSCNVLSLDLRESLIHQEDTIVFSLIERAKPP
LNSPAPESRCLDSGFSGLTEFFVRETEIIQAKVGRYEPENPFLENIHPSVPT
HKYPSALPKALSVINQINDIYFKELPLFVKPGDDGNVPSYTAASDLAQLASER
IHYGFVKEVFRDAPQYEPALRQBEALMKLLTPKEVEYMKVQKAEFTGGE
VAFNSGYGDESKKTKVDPLLASRIYGEWLFIPLNVEVEYLLRLD"
ORIGIN

Alignment Scores:
Pred. No.:      1..73e-70      Length:      1006
Score:          818.00      Matches:      162
Percent Similarity: 72.83%      Conservative: 21
Best Local Similarity: 61.13%      Mismatches: 68
Query Match:      61.09%      Indels:      4
DB:              6      Gaps:      2

US-10-624-061-16 (1-261) x AX343933 (1-1006)

QY      1 MetAlaLysAlaAlaGluClnSerProAspSerGly-----AsnValTyThrLeuAla 18
Db      64 ATGGCAAGAGATTTCGAATTCGGATTTCGGGTTCTGCTTCCAAATGACTAGCTTCGAT 123

```

```

QY      19 SerValArgGluAspLeuValArgGlnGlnAspThrIleIleTyGlyLeuIleGluArg 38
Db      124 TTAATCAGAGATCGTTGATTAGGCAAGAGACACCATCGTCTCAGCTTGTATCAGAGAGA 183
QY      39 AlaLysPheProSerAsnSerHisThrTyAspGluLysTyAlaGlnIleGlnGlyPhe 58
Db      184 GCTAAGTTTCCACTCAATCTCCTGCTTCGAGGAATCTCGTTGCTAGATTCTGGAAGT 243
QY      59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78
Db      244 TTCTCTTCTCTCACTGAGTTTTCGTCAGAGAGACAGAAATCATCCAAGCTAAGTAGGA 303
QY      79 ArgTyLysAsnProGluGlnAsnAlaPhePheProGluAsnLeuProProSerIleVal 98
Db      304 AGATATGATACCCGGAAGAGAAATCCTTCTTCCTTGAGAAATCTCCTCAGCTGGTTT 363
QY      99 ProSerTySerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118
Db      364 CCTACGCAAAATATCCATCGGCTTTGCACCCCTAAGGCTCTATCTGTTAAACATTACCAA 423
QY      119 SerIleTrpLysMetTyThrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138
Db      424 CAAATCTGGGATATTTACTTTAAAGAAATTCCTTTCTTTGTCAAACCTGGCGATGAT 483
QY      139 GlyAsnTyAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158
Db      484 GGCAACTATCCATCACTGCTGCTAGTATCTCGCTGTTTACAAGCTCTTTCGAGAAAGG 543
QY      159 IleHisTyGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyGlu 178
Db      544 ATTCACTACGGTAAATTTCTAGCTGAGTCAAAATTCAGAGATGCTCCCAAGATTACGAG 603
QY      179 ProLeuileArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198
Db      604 CTGCGATTCGGCTCAGATAGAGAGGCTTTGATGAAGCTCTTGACGTTTGAGAAAGTA 663
QY      199 GluGluThrValArgLysArgValGluLysLysAlaValPheGlyGlnGluValAsn 218
Db      664 GAAGAAATGTTTAAAGAGAGAGTGCAGAGAAAGAGAGAAACGTTTGGCAAGAAAGTAAA 723
QY      219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236
Db      724 TTCAACTCTGGCTATGGCGATCAGAGTAAGAGAGATATAAGTGCATCCATGCTTGC 783
QY      237 SerSerLeuTyLysAsnTrpValPheProLeuThrLysGluValGlnValGluTyLys 256
Db      784 TCTCGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTT 843
QY      257 LeuArgArgLeuAsp 261
Db      844 CTACGTCGTCGAT 858

RESULT 8
AX464575      1006 bp      DNA      linear      PAT 16-JUL-2002
LOCUS
DEFINITION
Sequence 3 from Patent WO020798.
ACCESSION
AX464575
VERSION
AX464575.1 GI:21899370
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
Boyes, D.C., Davis, K.R., Woessner, J.P., Goriach, J., Hamilton, C.M.,
Hoffman, N.E., Kloti, A.S., Zayed, A., Ascenzi, R.A., Allen, K.,
Mulpuri, R. and Kjemtrup, S.
Methods and compositions for the modulation of chorismate synthase
and chorismate mutase expression or activity in plants
Patent: WO 0202798-A 3 10-JAN-2002;
Paradigm Genetics Inc. (US)
FEATURES
Location/Qualifiers

```

```

source      1..1006
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Alignment Scores:
Pred. No.:      1,73e-70      Length:      1006
Score:          818.00      Matches:      162
Percent Similarity: 72.83%      Conservative: 31
Best Local Similarity: 61.13%      Mismatches: 68
Query Match:    61.09%      Indels:      4
DB:             6          Gaps:      2

US-10-624-061-16 (1-261) x AX464575 (1-1006)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18
Db 64 ATGGCAAGAGTCTTCGAATCGGATTCGGGTCTGGTTGTTCCAAATGACTGAGTCTTGAC 123
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuLeuGluArg 38
Db 124 TTAATCAGAGAATCGTGTGATAGGCAAGAGACACCATCGTCTTCAGCTTGATCGAGAGA 183
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnLeuGlnGlyPhe 58
Db 184 GCTPAGTTTCCACTCAATCTCTCGTTCGAGGAATCTCGTGTCTAGATTCTCGAAGT 243
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78
Db 244 TTCTCTTCTCACTCAGTGTTCGTCAGAGACAGACAATCATCCAACTAAGGTAGGA 303
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProSerIleVal 98
Db 304 AGATATGAATACCCGGAAGAGAACTCTTCTTCCTTGAGAAACATCTCTCACTCGGTTTTT 363
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLys 118
Db 364 CCTACGACAAATATCCATCCGCTTTGACCCCTAAGGCTCTATCTGTTAACATTAAACAA 423
QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138
Db 424 CAATCTGGGATATTACTTTAAAGAAATGCTTCTCTTGTGTCMAAACCCTGGCGATGAT 483
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158
Db 484 GGCACATATCCATCACTGCTGCTAGTATCTCGCCTGTTTACAGCTCTTTCGAGAGG 543
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178
Db 544 ATTCACTACGGTAAATTTGATGCTGAGGTCAAATTCAGAGATGCTCCACAAGATTACGAG 603
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198
Db 604 CCTGCGGATTCGCGCTCAGGATAGAGAGGCTTTGATGAGCTGTTGACGTTTGAGAAAGTA 663
QY 199 GluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsn 218
Db 664 GAAGAAATGTTAAGAAGAGAGTGCAGAGAAAGCAAGCAAGTTTCGACAAAGATTAAGA 723
QY 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236
Db 724 TTCAACTCTGGCTATGCGGATGAGAGTAAGAAGATATAAAGTGGATCCATTGCTTTGCC 783
QY 237 SerSerLeuTyrLysAsnTyrValIleProLeuThrLysGluValGlnValGluTyrLeu 256
Db 784 TCTCGCATCTACGGGAATGCTTTATCCCTCTCACTAAGCTCGTTGAGGTGAGTATCTT 843
QY 257 LeuArgArgLeuAsp 261
Db 844 CTACGTCGTCGAT 858

RESULT 9

```

```

AX464581/c
LOCUS      AX464581      1006 bp      DNA      linear      PAT 16-JUL-2002
DEFINITION Sequence 9 from Patent WO0202798.
ACCESSION  AX464581
VERSION     AX464581.1  GI:21899376
KEYWORDS
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE   1
AUTHORS     Boyes D.C., Davis K.R., Woessner J.P., Gorlach J., Hamilton C.M.,
            Hoffmann N.E., Klotz A.S., Zayed A., Ascenzi R.A., Allen K.,
            Mulpuri R. and Kjemtrup S.
TITLE       Methods and compositions for the modulation of chorismate synthase
            and chorismate mutase expression or activity in plants
JOURNAL     Patent: WO 0202798-A 9 10-JAN-2002;
            Paradigm Genetics Inc. (US)
FEATURES
            source
            1..1006
            /organism="Arabidopsis thaliana"
            /mol_type="unassigned DNA"
            /db_xref="taxon:3702"

ORIGIN

Alignment Scores:
Pred. No.:      1,73e-70      Length:      1006
Score:          818.00      Matches:      162
Percent Similarity: 72.83%      Conservative: 31
Best Local Similarity: 61.13%      Mismatches: 68
Query Match:    61.09%      Indels:      4
DB:             6          Gaps:      2

US-10-624-061-16 (1-261) x AX464581 (1-1006)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18
Db 943 ATGGCAAGAGTCTTCGAATCGGATTCGGGTCTGGTTGTTCCAAATGACTGAGTCTTGAC 884
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuLeuGluArg 38
Db 883 TTAATCAGAGAATCGTGTGATAGGCAAGAGACACCATCGTCTTCAGCTTGATCGAGAGA 824
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnLeuGlnGlyPhe 58
Db 823 GCTAAGTTTCCACTCAATCTCTCGTTCGAGGAATCTCGTGTCTAGATTCTGGAAGT 764
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78
Db 763 TTCTCTTCTCACTGAGTTTTTCTCTCAGAGACAGAGAAATCATCCAAAGCTTAAAGTAGGA 704
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleVal 98
Db 703 AGATATGAATACCCGGAAGAGAACTCTTCTCTTTCGAGAACTCTCTCACTCGGTTTTT 644
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLys 118
Db 643 CCTACGCAAAATATCCATCCGCTTTGCACCCCTAAGGCTCTATCTGTTAACATTAAACAA 584
QY 119 SerIleTyrLysMetTyrPheLysLeuLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138
Db 583 CAATCTGGGATATTACTTTAAGAATTTGTTCTTTGTTGTTGTCAAACCTGGCGATGAT 524
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158
Db 523 GGCACATATCCATCAACTGCTGCTAGTATCTCGCTGTTTACAAGCTCTTTCGAGAGG 464
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178
Db 463 ATTCACTACGGTAAATTTGATGCTGAGGTCAAATTCAGAGATGCTCCACAAGATTACGAG 404
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198

```

```

Db 403 CTGCGATTCGGCTCAGGATAGAGAGCTTTGATGAAGCTGTTGACGTTTGAGAAAAGTA 344
QY 199 GluGluThrValArgLysArgValGluLysLysAlaValPheGlyGlnGluValAsn 218
Db 343 GAAGAATGCTTAAAGAGAGAGTGCAGAGAAGAGCAAGAGCTTTGGACAGAGTAATA 284
QY 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236
Db 283 TTCAACTCTGGCTATCGGCATGAGAGTAAGAAGAGTATAAAGTGGATCCATTCGTTGCC 224
QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluLeu 256
Db 223 TCTCGCAATCTCGGGATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTGAGTAICTT 164
QY 257 LeuArgArgLeuAsp 261
Db 163 CTACGTCGTCGAT 149

```

```

RESULT 10
L47355
LOCUS
DEFINITION
Arabidopsis thaliana chorismate mutase mRNA, linear PLN 13-AUG-1999
ACCESSION
L47355
VERSION
L47355.1 GI:5732015
KEYWORDS
chorismate mutase; mutase.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1006)
Eberhard J., Ehrler T.T., Epple, P., Felix, G., Raesecke, H.R.,
Amrhein, N. and Schmid, J.
Cytosolic and plastidic chorismate mutase isozymes from Arabidopsis
thaliana: molecular characterization and enzymatic properties
Plant J. 10 (5), 815-821 (1996)
97111372
PUBMED
8953244
COMMENT
GSDS:S:46595
FEATURES
source
1..1006
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

polya_site
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

polya_site
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

polya_site
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

polya_site
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

polya_site
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

polya_site
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

polya_site
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

polya_site
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

polya_site
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

polya_site
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

polya_site
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="cAtCM2"
64..861
/EC_number="5.4.99.5"
/note="putative"
/codon_start=1
/product="chorismate mutase"
/db_xref="GI:5732015"
/translation="MARVFESDSGSGSNVLSDLIRESLIROEDTIVFSLIERAKFP
LNSAFESRCLDSGSFSSLTFFVRETEIIQAKVGRVEYEPENPFLENIPHSVPT
HKYFSALHPKALSNVINKQIWDIYFKELLPUFVRPDDGNYPSTAASDLACLQALSR
IHYGKFAVFRPAQDYPFAIRAQDEALMKLLTFKVEEMVKRVQKKQVETFGQE
VKFNSGYGDESKKYKVDPLLASRIYGEWLIPTLKUVEVYLLRLD"
1006
/note="putative"

```

```

Db 64 ATGGCAAGAGTCTTCCGAATCGGATTCGGGTTCGTTGTTCCAAATGTAAGTCTTGCAC 123
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleIleIleIleIleGluArg 38
Db 124 TTAATCAGAGATCGTTGATTAGGCAAGACACATCGTCTTCAGCTTGTATCGAGAGA 183
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnIlePhe 58
Db 184 GCTAAGTTTCCACTCAATTCCTGCTTCGTTTCGAGGAATCTCGTTGTCTAGATTCTGAA 243
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78
Db 244 TTCTCTTCTCTCACTGAGTTTTCGTGAGAGACAGAAATCATCCAGCTTAGGTAGGA 303
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleVal 98
Db 304 AGATATGATACCCGGAAGAGAATCTTTCTTCTTCTTGTGAAACATTCCTCCTCGGTTT 363
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118
Db 364 CCTACGCCAACAATATCCATCGGCTTTGCACCTTAAGGCTCTATCTGTAAACATTAACAA 423
QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138
Db 424 CAATCTGGGATATTACTTTAAAGAATTGCTTCTCTTGTTCACCACTGGCGATGAT 483
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArg 158
Db 484 GGCACTATCCATCAACTGCTCTAGTATCTCGCTGTTCACAGCTCTTTCGAGAGG 543
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGluAspTyrGlu 178
Db 544 ATTCATAGCTGAATTTGTAGCTGAGGTCAAATTCAGAGATGCTCCACAGATTAACGAG 603
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198
Db 604 CTGCGATTCGGGCTCAGGATAGAGAGCTTTGATGAAGCTGTTGACGCTTTGAGAGAATA 663
QY 199 GluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsn 218
Db 664 GAAGAAATGCTTAAAGAGAGAGTGCAGAGAAGAGCAAGAGCTTTCGACAGAGTAATA 723
QY 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236
Db 724 TTCAACTCTGCTATGGCGATGAGAGTAAGAAGAGTATAAAGTGGATCCATTCGTTGCC 783
QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluLeu 256
Db 784 TCTCGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTGAGTATCTT 843
QY 257 LeuArgArgLeuAsp 261
Db 844 CTACGTCGTCGAT 858

```

## RESULT 11

AY065238

LOCUS

DEFINITION

Arabidopsis thaliana putative chorismate mutase CM2 (At5g10870)

mRNA, complete cds.

ACCESSION

AY065238

VERSION

AY065238.1

KEYWORDS

FLI CDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1055)

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Ban, J., Chung, M.K.,

Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,

Carrinchi, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,

Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,

PLN 18-SEP-2002

linear

mRNA

1055 bp.

putative chorismate mutase CM2 (At5g10870)

Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Arabidopsis Full Length cDNA Clones

# TITLE JOURNAL REFERENCE

## AUTHORS

Unpublished  
2 (bases 1 to 1055)

Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (03-DEC-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.

## TITLE JOURNAL

## COMMENT

The Salk, Stanford, PGENC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGENC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGENC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

## FEATURES

## source

1..1055  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="RAFL09-41-E09 (R18739)"  
/ecotype="Columbia"  
/note="This clone is in a modified pBluescript vector (FIC-1) as a BamHI/XhoI insert."

## gene

1..1055  
/gene="At5g10870"

## 5'UTR

1..98  
/gene="At5g10870"

## misc\_difference

2  
/gene="At5g10870"  
/gene="At5g10870"  
/note="compared to genomic sequence"  
/replace="t"

## CDS

99..896  
/gene="At5g10870"  
/codon\_start=1  
/evidence=experimental  
/product="putative chorismate mutase CM2"  
/protein\_id="AA138714.1"  
/db\_xref="GI:17529008"  
/translation="MARVFESDGGSCNVSLDLRLRESLRQEDTIVFSLIERAKFP  
LNSPAPFESCLDSGFSLSLTFVRETEIIOAKVGRYEPENPFLENIPHSVPPT  
HKYPSALHPKALSVNINQIWDYIFKELPLFVKGDDGNYPSAAADLQALSR  
IHYGKFAVEKFRDAPDYEPARAODREALMLLTPEKVEEMVKRVQKKAETTFQE  
VKNSGGVDSKSKYKVDPLLASRIYGEMLIPTKLVEVEYLLRLD"  
897..1055  
/gene="At5g10870"

## 3'UTR

897..1055  
/gene="At5g10870"

## ORIGIN

## Alignment Scores:

Pred. No.:	1..83e-70	Length:	1055
Score:	818.00	Matches:	162
Percent Similarity:	72.83%	Conservative:	31
Best Local Similarity:	61.13%	Mismatches:	68
Query Match:	61.09%	Indels:	4
DB:	8	Gaps:	2
US-10-624-061-16 (1-261) x AY065238 (1-1055)			
QY	1	MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla	18
Db	99	ATGGCAAGAGTCTTCGAATCGGATTTCGGTTCCTGTTGTTCCAAATGACTAGTCTTGAC	158
QY	19	SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArg	38
Db	159	TTAAATCAGAGATCGTTGATTAGGCAAGAAGACACCATCGTCTTCAGGTTCATCAGAGA	218
QY	39	AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe	58
Db	219	GCTAAGTTTCCATCAATCTCCTGCTTCGAGGAATCTCGTTGTCTAGATTCTGGAAGT	278
QY	59	CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly	78
Db	279	TTCTCTTCTCTCAGTGGTTCGTCAGAGACAGAGAAATCATCCAAAGCTAAGTAGGA	338
QY	79	ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleVal	98
Db	339	AGATATGATATACCCGAGAGAGAAATCTTCTTCTTGAGAACATTCCTCACTCGGTTTT	398
QY	99	ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys	118
Db	399	CCTACGCACAAATATCCATCGGCTTTGCACCTAAGGCTCTATCTGTTAAACATTAACA	458
QY	119	SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp	138
Db	459	CAAAATCTGGATATTACTTTAAAGAAATGCTTCTCTTTGTTGTCAAACCTGGCGATG	518
QY	139	GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg	158
Db	519	GGCAACTATCCATCACTGCTGCTAGTATCTCGCTGTTTACAGCTCTTTCGAGAAGG	578
QY	159	IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu	178
Db	579	ATTCACTACGGTAAATTTGTAGCTGAGGTCAAATTCAGAGATGCTCCCAAGATTACG	638
QY	179	ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal	198
Db	639	CTTCGATTCGGCTCAGATAGAGAGGCTTTGATGAAGCTGTTGACGTTTGAGNAAGTA	698
QY	199	GluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsn	218
Db	699	GAAGAATATGTTAGAGAGAGATGCGAAGAAGCAGAAACGTTTGGACAAGAAATAAA	758
QY	219	LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla	236
Db	759	TTCAACTCTGGCTATGGCGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGATCT	818
QY	237	SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu	256
Db	819	TCTCGCATCTACGGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGATCTCT	878
QY	257	LeuArgArgLeuAsp	261
Db	879	CTACGCTGCTCGAT	893
RESULT 12			
AR404683			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
Unknown.			
Sequence 5 from patent US 6627798.			
525 bp DNA linear PAT 18-DEC-2003			

ORGANISM	Unknown.	JOURNAL	same, and methods of use
REFERENCE	Unclassified.	PATENT	WO 0216655-A 737 28-FEB-2002;
AUTHORS	1 (bases 1 to 525)	THE	Scipps Research Institute (US) ; Syngenta Participations AG
TITLE	Falco, S.C., Farnoud, O.O. and Lee, J.-M.	(CH)	
JOURNAL	Aromatic amino acid biosynthetic enzymes	FEATURES	Location/Qualifiers
PATENT	US 6627798-A 5 30-SEP-2003;	source	1..1005
FEATURES	Location/Qualifiers	ORIGIN	/db_xref="taxon:3702"
source	1..525		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:	6,41e-66	Length:	1005
Pred. No.:	767.50	Score:	525
Percent Similarity:	95.29%	Matches:	162
Best Local Similarity:	95.29%	Conservative:	0
Query Match:	57.32%	Mismatches:	7
DB:	6	Indels:	4
		Gaps:	0
US-10-624-061-16 (1-261) x AR404683 (1-525)			
QY	1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValThrLeuAlaSerVal	20	
DB	16 ATGGCCNAAGCAGCAGACAAAGTCTGATCTCGGAATGTGTACACCGTAGCTTCTGTG	75	
QY	21 ArgGluAspLeuValArgGlnGluAspThrIleLeuThrGlyLeuIleGluArgAlaLys	40	
DB	76 AGAGAGGATTGGTTAGGCAAGAGGATACCATCTATTATGGTCTCATTTGAGAGAGCCAA	135	
QY	41 PheProSerAsnSerHisThrTyrosPheGluGlnValThrAlaGlnIleGlnGlyPheCysGly	60	
DB	136 TTCCTTAGCAATCTCACACCTATGATGAAGAAAGTATGCTCAATCCAGGGTTTTGTGGC	195	
QY	61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr	80	
DB	196 TCATTGGTGGATTTGTTGTTAAGAAATACAGAGGCCATTCAGCTAAGGCTGAAGATAC	255	
QY	81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProSerIleValProSer	100	
DB	256 AAAACCCCTGAAGAAACCCCTCTTCCAGAAATTTTACCACCATCAATTTGTGCCATCT	315	
QY	101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLysSerIle	120	
DB	316 TACTCTCTCAACAGCTTTTGTGATCTCTGCTGCTTCAATTAACATAAACAAGTCACT	375	
QY	121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn	140	
DB	376 GGGAAATATGTTTCAAGAGATTACTTCCATTGCTTGTCTTCCGGTGTGATGGNAAC	435	
QY	141 TyrAlaGln-ThrAlaAlaAspLeuSerLeuLeuGlnSerIleSerArgArgIleHi	160	
DB	436 TATGGCAAACTGCAGCTAATGACCT-TCATATTGCAGTC-ATCTCTAGAGGAT-CA	492	
QY	160 sTyrGlyLysPheValAlaGluValLys	169	
DB	493 CTATGAAAGTTGTAGCTGANGGAAA	520	
RESULT 13			
LOCUS	AX506042	1005 bp	DNA linear PAT 27-SEP-2002
DEFINITION	Sequence 737 from Patent WO0216655.		
ACCESSION	AX506042		
VERSION	AX506042.1	GI:23387279	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE	1		
AUTHORS	Harper J.F., Kreps, J., Wang, X. and Zhu, T.		
TITLE	Stress-regulated genes of plants, transgenic plants containing		



```
RESULT 14
AR236633
LOCUS AR236633 1207 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 2 from patent US 6465217.
ACCESSION AR236633
VERSION AR236633.1 GI:27280719
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1207)
AUTHORS Boyes,D.C., Davis,K.R., Woessner,J.P., Goriach,J., Hamilton,C.M.,
Hoffman,N.E., Kloti,A.S., Zayed,A. and Ascenzi,R.A.
TITLE Methods and compositions for the modulation of chorismate synthase
and chorismate mutase expression or activity in plants
JOURNAL Patent: US 6465217-A 2 13-OCT-2002;
FEATURES Location/Qualifiers
source 1..1207
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,74e-58 Length: 1207
Score: 696.00 Matches: 131
Percent Similarity: 68.08% Conservative: 46
Best Local Similarity: 50.38% Mismatches: 75
Query Match: 51.98% Indels: 8
Gaps: 6
DB: 1

US-10-624-061-16 (1-261) x AR236633 (1-1207)
QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
DB 235 GATCAGAGTGCAGATTGACTCTTGAAGGTATTAGAACTCTTTGATCGTCAAGAGGAC 294
QY 30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49
DB 295 AGCATTATATTGGGCTATTGGAGAGAGCCAGTACTGTACAACTGCTGATCTATGAT 354
QY 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69
DB 355 CCTACTGCTTTGCAGATGGTGTTCATGCTTCTTGGTTGAGTACATGGTTAAAGGC 414
QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluAsnAlaPhePhe 89
DB 415 ACTGAGAAGCTTCACGCTAAGTGTGGTAGGTAAAGAGTCCCTGATGAACATCCTTTCTTC 474
QY 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
DB 475 CCTGATGATCTACAGAGCTATGTTGCTCTCTCTCAGTACCCAAAGGTGTTCATTTT 534
QY 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeu 129
DB 535 GCTGCTGATTCGATAACATAACAGAGATATGGAACATGATCTTACAGACCTTGT 594
QY 130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAsp 149
DB 595 CCAAGACTTCTGAAGAAGCCGATGATGTAACTACCGCTCAACACGCTGCTGTGACGCT 654
QY 150 SerLeuLeuGlnSerIleSerArgGileHisTyrGlyLysPheValAlaGluValLys 169
DB 655 ATCTGCTTCAGTGTCTCTCAAGAGATATCCATTACGGTAATTTGTTGAGAGCTAA 714
QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189
DB 715 TTTCAAGCCCTCACCGAAGCATACGAGTCCGCCATCAAGACCAAGATAAGATCGACTG 774
QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
DB 775 ATGATATGCTGACATTCCTCCGACTGTGGAAGATGCCGATAAAGAGAGATTGATGAA 834
QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAsnAspAsnGluAsnArg 229

835 ACCGACATACGGCAAGAGTGAAGTTGGGATGGAGGCAAGAGAGAGAGAGAA 894
230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
895 GAAGGGAATGAATCTCATGTTTACAAAATCATGCTCCGATCTTAGTTGGTGAATATGGA 954
242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuValArgLeuAsp 261
955 GATTGGATCATGCTCTTTACAAAAGAGTTCAAGTGGAGTACTTGTCTCAGAAGACTGGAC 1014

RESULT 15
AR236639/c
LOCUS AR236639 1207 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 8 from patent US 6465217.
ACCESSION AR236639
VERSION AR236639.1 GI:27280725
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1207)
AUTHORS Boyes,D.C., Davis,K.R., Woessner,J.P., Goriach,J., Hamilton,C.M.,
Hoffman,N.E., Kloti,A.S., Zayed,A. and Ascenzi,R.A.
TITLE Methods and compositions for the modulation of chorismate synthase
and chorismate mutase expression or activity in plants
JOURNAL Patent: US 6465217-A 8 15-OCT-2002;
FEATURES Location/Qualifiers
source 1..1207
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,74e-58 Length: 1207
Score: 696.00 Matches: 131
Percent Similarity: 68.08% Conservative: 46
Best Local Similarity: 50.38% Mismatches: 75
Query Match: 51.98% Indels: 8
Gaps: 6
DB: 1

US-10-624-061-16 (1-261) x AR236639 (1-1207)
QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
DB 973 GATCAGAGTGCAGATTGACTCTTGAAGGTATTAGAACTCTTTGATCGTCAAGAGGAC 914
QY 30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49
DB 913 AGCATTATATTGGGCTATTGGAGAGAGCCAGTACTGTACATGCTGATCTATGAT 854
QY 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69
DB 853 CCTACTGCTTTGACATGGATGCTTCAATGGTCTTTGGTTGAGTACATGTTAAAGGC 794
QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluAsnAlaPhePhe 89
DB 793 ACTGAGAAGCTTCACGCTAAGTGTGGTAGGTAAAGAGTCCCTGATGAACATCCTTTCTTC 734
QY 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
DB 733 CCTGATGATCTACAGAGCTATGTTGCTCTCTCTCAGTACCCAAAGGTGTTCATTTT 674
QY 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeu 129
DB 673 GCTGCTGATTCGATAACATAACAGAGATATGGAACATGATCTTCTCAGAGACTGTT 614
QY 130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAsp 149
DB 613 CCAAGACTTCTGAAGAAGCCGATGATGTAACTACCGCTCAACAGCTGCTGTGACGCT 554
QY 150 SerLeuLeuGlnSerIleSerArgGileHisTyrGlyLysPheValAlaGluValLys 169
```

```

Db      553 ATCTGCTTACGTCTCTCAAAGAGAATCCATTACGGTAAATTTGTCAGAAAGCTAAA 494
Qy      170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189
Db      493 TTTCAGGCTCACCCGAGCATACGAGTCGCCCATCAAAGCACAGATAGATCGACTG 434
Qy      190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
Db      433 ATGGATATGCTCACATTCCTCCGACTCTGGAAGATGCGATAAAGAGAGAGATTGAGATGAAA 374
Qy      210 AlaValValPheGlyGluGluValAsnLeuAsnSerAspAspAsnAspAsnGluAsnArg 229
Db      373 ACCCGAACATACGGGCAAGAGTGAAGTTGGGATGGAGGAGAAAGAGAAGAGAGAGAA 314
Qy      230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
Db      313 GAAGGGAAATGAATCTCATGTTTACAAATCAGTCCGATCTTAGTTGGTGACTTATATGGA 254
Qy      242 AsnTyrValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261
Db      253 GATTGGATCATGCCCTTTAAACAAAGAGGTTCAAGTGGAGTACTTCTCAGAAAGACTGGAC 194

```

Search completed: September 25, 2004, 04:17:47  
 Job time : 4199 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 25, 2004, 02:54:52 ; Search time 433 Seconds  
(without alignments)

2560.692 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKAAEQSPDGNVYTLASV.....NWVPLTKEVQVEYLLRLRD 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q=/cgm2\_1/USPTO.spool/US10624061/runat\_24092004.092406.16902/app\_query.fasta\_1.455

-DB=N Geneseq\_29Jan04 -OFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10624061@cgm2\_1.470 @runat\_24092004.092406.16902 -NCPUS=6 -ICPU=3

-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Geneseq\_29Jan04.\*

2: Geneseqn1990s.\*

3: Geneseqn1990s.\*

4: Geneseqn2000s.\*

5: Geneseqn2001as.\*

6: Geneseqn2001bs.\*

7: Geneseqn2002as.\*

8: Geneseqn2003as.\*

9: Geneseqn2003bs.\*

10: Geneseqn2003cs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1339	100.0	1020	7	ABX93058 Soybean c
2	818	61.1	798	6	ABZ13167 Arabidops
3	818	61.1	993	6	ABA99614 Construct
4	818	61.1	1006	6	ABA99613 A. thalia
5	818	61.1	1006	6	ABA91400 Arabidops
6	818	61.1	1006	6	ABA91394 Arabidops
7	767.5	57.3	525	7	ABX93053 Soybean c
8	698	52.1	1005	6	ABZ12932 Arabidops

c	9	696	52.0	1207	6	ABA91393 Arabidops
10	696	52.0	1207	6	ABA91399 Arabidops	
11	685.5	51.2	1231	7	ABX93057 Rice chor	
12	677.5	50.6	1274	3	AAC33324 Arabidops	
13	670.5	50.1	1223	7	ABX93056 Corn chor	
14	654.5	48.9	1217	6	ABA91395 Arabidops	
15	654.5	48.9	1217	6	ABA91401 Arabidops	
16	602	45.0	780	7	ABX93059 Wheat cho	
17	587	43.8	5176	6	ABA91397 Arabidops	
18	549.5	41.0	5099	6	AAD29643 Arabidops	
19	549.5	41.0	5099	6	ABA91396 Arabidops	
20	522	39.0	579	7	ABX93054 Wheat cho	
21	423	31.6	843	3	AAC81949 H. polyo	
22	423	31.6	1655	3	AAC81950 H. polyo	
23	402.5	30.1	869	7	AAL50202 M sterili	
24	371.5	27.7	2920	7	ABT17891 Aspergill	
25	371.5	27.7	2987	7	ABT19705 Aspergill	
26	312.5	23.3	658	3	AAF14469 Aspergill	
27	267	19.9	287	6	ABL72637 Corn tass	
28	248	18.5	258	6	ABL71526 Corn tass	
29	144.5	10.8	560	4	AAH44222 Physcomit	
30	144.5	10.8	601	7	ABX93052 Rice chor	
31	130	9.7	552	7	ABX93051 Corn chor	
32	121.5	9.1	1379	6	ABA91794 Yeast mit	
33	113	8.4	507	3	AAF07985 Fusarium	
34	105	7.8	1080	7	ABX05983 S. pneumo	
35	105	7.8	19250	2	AAV52168 Streptoco	
36	105	7.8	110000	4	ABE54544_02 Continuation (3 of	
37	103	7.7	2346	4	AAS53535 Haemophil	
38	103	7.7	2346	7	ACA34520 Prokaryot	
39	103	7.7	110000	2	AAT42063_17 Continuation (18 o	
40	100.5	7.5	2385	9	ADD30303 Plant yie	
41	96.5	7.2	2043	6	ABZ13532 Arabidops	
42	96.5	7.2	2043	8	ABZ23149 Enviroame	
43	96	7.2	1380	4	AAH53580 S. epider	
44	96	7.2	1380	4	AAH52926 S. epider	
45	96	7.2	1395	6	ABN92401 Staphyloc	

#### ALIGNMENTS

RESULT 1

ABX93058

ID ABX93058 standard; cDNA; 1020 BP.

XX AC ABX93058;

XX DT 22-MAY-2003 (first entry)

XX DE Soybean chorismate mutase cDNA.

XX KW Soybean; plant; gene; ss; chorismate mutase; aromatic amino acid; diet;  
XX KW biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;  
XX KW phenylalanine; prephenate dehydrogenase; herbicide; food crop.

XX OS Glycine max.

XX FH Key Location/Qualifiers

XX FT CDS 23..808

XX FT /tag= a

XX FT /product= "Chorismate mutase"

XX PN US2002184658-A1.

XX PD 05-DEC-2002.

XX PF 03-DEC-1999; 99US-00454279.

XX PR 04-DEC-1998; 98US-0110845P.

XX PA (FALC/) FALCO S C.

XX PA (FAMO/) FAMODU O O.

XX PA (LEE/) LEE J.

XX Falco SC, Famodu OO, Lee J;  
 XX WPI; 2003-328651/31.  
 DR P-PSDB; ABU08094.  
 XX  
 PT New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
 PT nucleic acids, useful for facilitating design and/or identifying  
 PT inhibitors of those enzymes that may be used as herbicides and for  
 PT producing antibodies.  
 XX  
 PS Claim 3; Page 22; 32pp; English.  
 XX  
 CC The invention discloses isolated polynucleotides encoding chorismate  
 CC mutase polypeptides. Aromatic amino acids must be included in the diets  
 CC of animals. In the aromatic amino acid biosynthetic pathway chorismate is  
 CC converted to anthranilate during tryptophan biosynthesis and is converted  
 CC to prephenate, the branch point for tyrosine and phenylalanine  
 CC biosynthesis. Chorismate mutase catalyses the conversion of chorismate to  
 CC prephenate. Also disclosed are methods for selection of an isolated  
 CC polynucleotide that affects the level of expression of a tyrosine  
 CC biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
 CC fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
 CC evaluating compounds for their ability to inhibit the activity of a  
 CC tyrosine biosynthetic enzyme. The polypeptides can be used produce  
 CC antibodies. Chorismate mutase and prephenate dehydrogenase are good  
 CC targets for herbicides that will not affect animals, and overexpression  
 CC of these enzymes may be used to increase the content of aromatic amino  
 CC acid in food crops. The polypeptides may also be used to design and/or  
 CC identify inhibitors of those enzymes that may be used as herbicides. The  
 CC nucleic acids may be used to create transgenic plants, as probes for the  
 CC genetic and physical mapping of the genes and as markers for traits  
 CC linked to those genes. The sequence presented is a soybean chorismate  
 CC mutase cDNA  
 XX  
 SQ Sequence 1020 BP; 322 A; 176 C; 223 G; 299 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1,32e-148 Length: 1020  
 Score: 1339.00 Matches: 261  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-624-061-16 (1-261) x ABX93058 (1-1020)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
 DB 23 ATGGCCAAAGCAGCAGAACAAAGCTCTGATCTGGGAATGTACACGCTAGCTTCGTG 82  
 QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLys 40  
 DB 83 AGAGAGGATTGGTTAGCAAGAGGATACCATCTATTTATGCTCTCATTTAGAGAGCCAA 142  
 QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
 DB 143 TTCCTAGCAATTCACACCATATGATGANAAGATGCTCAATCCAGGGTTTGTGGC 202  
 QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
 DB 203 TCATTGGTGGATTTGTTTAAAGATACAGAGCCATTCAGCTAGGCTGGAGATAC 262  
 QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100  
 DB 263 AAAAACCCCTGAAGAAACGCCCTCTCTCCAGAAAATTTACCATCAATTTGTCCTATC 322  
 QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLysSerIle 120  
 DB 323 TACTCTCTCAACACAGTTTTCATCTCTGCTGCTCAATTACATTAACACAGTCCATC 382  
 QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsn 140

DB 383 TGGAAATGTATTTCAAAGAGTTACTTCCATTGCTTGGTGTATTCGGGTGATGCGCAAC 442  
 QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHis 160  
 DB 443 TATGCGCAAACTGCAGCTTAATGACCTTTTCATTATTGACGTCCATCTCTAGAAGGATTAC 502  
 QY 161 TyrGlyPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
 DB 503 TATGGAAGTTTGTAGCTAGGTGAAATTCAGGGATGCTCTCAAGACTACGAGCTTTA 562  
 QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200  
 DB 563 ATTCGAGCTAAGATTAAGAGGATTTGATGAATTTTGACATTTACAGCGTTGAGAG 622  
 QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220  
 DB 623 ACGGTGAGGAAGAGAGTTGAAAAGAGGCTGTGTGTTTGGGCAGGAGTGAATCTTAAC 682  
 QY 221 SerAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240  
 DB 683 AGTCATGACATGACATTAAGAAACCGTAATTTGATCCATCAGTGGCTTCTAGCTTGTAC 742  
 QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeu 260  
 DB 743 AAAAATTTGGTGTATCTCTCACCAGGAGGTTTCAAGTTGAGTCTCTTGGCGGTCTA 802  
 QY 261 Asp 261  
 DB 803 GAC 805  
 RESULT 2  
 ABZ13167  
 ID ABZ13167 standard; DNA; 798 BP.  
 XX  
 AC ABZ13167;  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 972.  
 XX  
 KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
 XX  
 CS Arabidopsis thaliana.  
 XX  
 FN WO200216655-A2.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 24-AUG-2001; 2001WO-US026685.  
 XX  
 PR 24-AUG-2000; 2000US-0227866P.  
 PR 26-JAN-2001; 2001US-0264647P.  
 PR 22-JUN-2001; 2001US-0300111P.  
 XX  
 PA (SCRI ) SCRIPPS RES INST.  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Harper JF, Krepis J, Wang X, Zhu T;  
 XX  
 WPI; 2002-304127/34.  
 XX  
 PT Identifying a stress condition to which a plant cell has been exposed and  
 PT producing plants with increased tolerance to these abiotic stresses.  
 XX  
 PS Claim 144; SEQ ID NO 972; 577pp + Sequence Listing; English.  
 CC The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising: (a) contacting nucleic acid  
 CC representative of expressed polynucleotides in the plant cell with an  
 CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants

CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office  
 XX  
 XX  
 SQ Sequence 798 BP; 223 A; 162 C; 180 G; 233 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 5.09e-87 Length: 798  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservative: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 6 Gaps: 2

US-10-624-061-16 (1-261) x ABZ13167 (1-798)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 Db 1 ATGGCAAGAGTCTCGAATCGGATTCGGGTTCGGTGTTCGAAGTACTGAGTCTTGAC 60  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleIleIleIleIleIleIle 38  
 Db 61 TTAATCAGAGATCGTTGATTAGGCAAGAACACCATCGTCTTCAGCTTGATCGAGAGA 120  
 QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 Db 121 GCTAAGTTTCCACTCAATCTCTGCTTTCGAGGAATCTCGTGTCTGATTCGGAAGT 180  
 QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 Db 181 TTCCTCTCTCACTAGTTTTTCGTCAGAGACAGAAATCATCAACCTGAAGTAGGA 240  
 QY 79 ArgTyrLysAsnProGluGluAlaPhePheProGluAsnLeuProSerIleVal 98  
 Db 241 AGATATGAATACCCGGAAGAGATCTCTTCTTCCTTGAGAACATCTCTCACTCGGTTTT 300  
 QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
 Db 301 CTTACGCACAAATATCCATCGGCTTTGACCCCTAAGCTCTATCTGTTAACATTAAACA 360  
 QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138  
 Db 361 CAAATCTGGGATATTACTTAAAGAAATTCCTTCCTTTGTCAAACTGGCGATGAT 420  
 QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158  
 Db 421 GGCACATATCCATCACTGCTGATGATCTCGCTGTTTACAGCTCTTTTCGAGAAGG 480  
 QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 Db 481 ATTCACCTACGGTAAATTTGTAGCTGAGGTCAAATTCAGAGATGCTCCACAAGATTACGAG 540  
 QY 179 ProIleuLeuArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198  
 Db 541 CCTGCGGATTCGCGCTCAGGATAGAGAGCGCTTTCATGAGCTGTGACGTTTGAGAAAGTA 600  
 QY 199 GluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsn 218  
 Db 601 GAAGAATGTTAAGAGAGAGTGCAGAGAAGACAGAAACGTTTCGACAGAGAGTAAAA 660  
 QY 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236  
 Db 661 TTCAACTCTGGCTATGCGGATGAGAGTAAAGAGAGTATAAAGTGCATTCATGCTTGGC 720  
 QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
 Db 721 TCTCGGATCTACCGGGAATGGCTTATCCCTCTCACTAAGCTGTTGAGGTGATGATCT 780  
 QY 257 LeuArgArgLeuAsp 261  
 Db 781 CTACGTCGTCCTCGAT 795

RESULT 3  
 ABA99614

ID ABA99614 standard; DNA; 993 BP.

XX AC ABA99614;

XX DT 17-MAY-2002 (first entry)

XX DE Construct PCR-script/AtCM-1TP-AtCM2 DNA.

XX KW Shikimate pathway; chorismate mutase; gene; prephenate dehydrogenase;  
 XX vitamin E; vitamin K; ubiquinone; fat-soluble antioxidant; transgenic;  
 KW feed; pharmaceutical; cosmetic; aromatic amino acid; salicylic acid;  
 KW folic acid; phenylpropanoid; flavanoid; plant; abiotic stress resistance;  
 KW tocopherol; tocotrienol; ds.

XX OS Arabidopsis thaliana.

XX OS Synthetic.

XX PN WO200200901-A1.

XX PD 03-JAN-2002.

XX PF 28-JUN-2001; 2001WO-EP007391.

XX PR 29-JUN-2000; 2000DE-01030647.

XX PR 21-DEC-2000; 2000DE-01064454.

XX PA (SUNG-) SUNGENE GMBH & CO KGAA.

XX PI Badur R, Geiger M, Kunze I, Sommer S;

XX DR WPI; 2002-164442/21.

XX DR P-PSDB; AAM49655.

XX PT Preparing fine chemicals, particularly Vitamins E and K, useful as  
 PT antioxidants e.g. in foods or medicine, by growing organisms with altered  
 PT shikimate biosynthesis pathway.

XX PS Claim 11; Page 62-63; 74pp; German.

XX CC This invention describes a novel method for the preparation of fine  
 CC chemicals by culturing organisms in which the shikimate pathway has been  
 CC altered relative to the wild type. The method involves a construct  
 CC containing a plastid transit peptide, a chorismate mutase and/or  
 CC prephenate dehydrogenase, linked to at least one regulatory sequence for  
 CC transcription or translation in plants. The method is used to produce  
 CC chemicals, particularly Vitamins E or K and/or ubiquinone, useful e.g. as  
 CC fat-soluble antioxidants (in feeds, pharmaceuticals and cosmetics), but  
 CC also aromatic amino acids, salicylic or folic acid derivatives,  
 CC phenylpropanoids, flavanoids etc., especially in transgenic plants. Also  
 CC plants with increased Vitamin E contents have improved resistance to  
 CC abiotic stress, e.g. frost and drought. Transgenic plants with an altered  
 CC shikimate pathway are useful as foods, fodder and in preparation of  
 CC processed foodstuffs. Transgenic plants with modified shikimate pathways  
 CC have increased content of desired chemicals, particularly tocopherols  
 CC and/or tocotrienols. This sequence encodes the construct PCR-script/AtCM-  
 CC 1TP-AtCM-2 which is composed of the Arabidopsis thaliana plastid  
 CC chorismate mutase-1 transit peptide and an Arabidopsis thaliana chorismate  
 CC mutase-2 fragment and is used for plastid localisation

SQ Sequence 993 BP; 259 A; 227 C; 210 G; 297 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.93e-87 Length: 993  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservative: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 6 Gaps: 2

US-10-624-061-16 (1-261) x ABA99614 (1-993)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 Db 196 ATGGCAAGAGCTTCGAATCGAATCGGTTCTGGTTCCCAATGACTAGTCTTGAC 255  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArg 38  
 Db 256 TTAATCAGAGAATCGTTGATTAGGCAAGAACACCATCGTCTTCAGCTTGATCAGAGA 315  
 QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 Db 316 GCTAAGTTTCCACTCAATCTCCCTGCTTCGAGGAATCTGCTGTAGATTCTGGAGT 375  
 QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 Db 376 TTCCTCTCTCACTGAGCTTTTCTGTCAGAGACAGAAATCATCCAAAGCTTAAGTAGGA 435  
 QY 79 ArgTyrLysAsnProGluGlnAsnAlaPhePheProGluAsnLeuProSerIleVal 98  
 Db 436 AGATATGAATACCCGGAGAGAAATCCITTCCTTGAGAAATCTCTCACTCGGTTT 495  
 QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
 Db 496 CCTACGCAAAATATCCATCGCTTTGCACCTTAAGGCTCTATCTGTTAACTAACA 555  
 QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138  
 Db 556 CAAATCTGGATATTACTTTAAAGAATGTCTCTTTGTTGCAACCTGGCGATGAT 615  
 QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArg 158  
 Db 616 GCAACTATCCATCACTGCTGCTAGTATCTCGCTCTTTACAGCTCTTCGAGAAGG 675  
 QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 Db 676 ATTCACTACGGTAAATTTGTAGCTGAGGTCAAATTCAGAGATGCTCCACAGATTACGAG 735  
 QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198  
 Db 736 CTGGATTCGGCTTCAGATAGAGAGCTTTGATGAAGCTGTTCAGCTTGAGAAGTA 795  
 QY 199 GluGluThrValArgLysArgValGluLysLysAlaValAlaPheGlyGlnGluValAsn 218  
 Db 796 GAAGAATGCTTTAAGACAGAGTGCAGAAAGACAGAAAGCTTTGGACAAAGTA 855  
 QY 219 LeuAsnSerAspAspAsnAsnGluAsnArgLysPhe-----AspProSerValAla 236  
 Db 856 TTCAACTCTGGCTATGGCGATGAGAGTAAGAGAGATATAAGTGCATCCATCTGCCC 915  
 QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
 Db 916 TCTCGCATCTACGGGAATGGTTATCTCTCTCACTAAGCTCGTTGAGTTGAGTATCTT 975  
 QY 257 LeuArgArgLeuAsp 261  
 Db 976 CTACGCTGCTCCGAT

## RESULT 4

IDA ABA99613 standard; DNA; 1006 BP.

XX ABA99613;

AC ABA99613;

XX 17-MAY-2002 (first entry)

DT

XX A. thaliana cytosolic chorismate mutase-2 DNA.

DE

XX Shikimate pathway; chorismate mutase; gene; prephenate dehydrogenase;

KW vitamin E; vitamin K; ubiquinone; fat-soluble antioxidant; transgenic;

KW feed; pharmaceutical; aromatic; aromatic amino acid; salicylic acid;

KW folic acid; phenylpropanoid; flavanoid; plant; abiotic stress resistance;

KW tocopherol; tocotrienol; ds.

XX

OS Arabidopsis thaliana.  
 XX Key Location/Qualifiers  
 FH 64..861  
 FT CDS  
 PT /tag= a  
 FT /product= "chorismate mutase-2"  
 XX  
 PN WO200200901-A1.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 28-JUN-2001; 2001WO-EP007391.  
 XX  
 PR 29-JUN-2000; 2000DE-01030647.  
 PR 21-DEC-2000; 2000DE-01064454.  
 XX  
 PA (SUNG-) SUNGENE GMBH & CO KGAA.

XX Badur R, Geiger M, Kunze I, Sommer S;  
 XX WPI; 2002-164442/21.  
 DR P-PSDB; AAM49654.  
 XX  
 XX Preparing fine chemicals, particularly Vitamins E and K, useful as  
 PT antioxidants e.g. in foods or medicine, by growing organisms with altered  
 PT shikimate biosynthesis pathway.  
 XX  
 PS Disclosure; Page 57-60; 74pp; German.

XX This invention describes a novel method for the preparation of fine  
 CC chemicals by culturing organisms in which the shikimate pathway has been  
 CC altered relative to the wild type. The method involves a construct  
 CC containing a plasmid transit peptide, a chorismate mutase and/or  
 CC prephenate dehydrogenase, linked to at least one regulatory sequence for  
 CC transcription or translation in plants. The method is used to produce  
 CC chemicals, particularly Vitamins E or K and/or ubiquinone, useful e.g. as  
 CC fat-soluble antioxidants (in feeds, pharmaceuticals and cosmetics), but  
 CC also aromatic amino acids, salicylic or folic acid derivatives,  
 CC phenylpropanoids, flavonoids etc., especially in transgenic plants. Also  
 CC plants with increased Vitamin E contents have improved resistance to  
 CC abiotic stress, e.g. frost and drought. Transgenic plants with an altered  
 CC shikimate pathway are useful as foods, fodder and in preparation of  
 CC processed foodstuffs. Transgenic plants with modified shikimate pathways  
 CC have increased content of desired chemicals, particularly tocopherols  
 CC and/or tocotrienols. This sequence encodes the Arabidopsis thaliana  
 CC cytosolic chorismate mutase-2 protein used in the method of the invention

XX Sequence 1006 BP; 290 A; 188 C; 226 G; 302 T; 0 U; 0 Other;

XX Alignment Scores:  
 Pred. No.: 7,06e-87 Length: 1006  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservative: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: Gaps: 2

US-10-624-061-16 (1-261) x ABA99613 (1-1006)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 Db 64 ATGGCAAGAGCTTCGAATCGGATTCGGTTCTGTTCCCAATGACTAGTCTTGAC 123  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArg 38  
 Db 124 TTAATCAGAGAATCGTTGATTAGGCAAGAACACCATCGTCTTCAGCTTGATCAGAGA 183  
 QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 Db 184 GCTAAGTTTCCACTCAATCTCTCTCTTCGAGGAATCTCGTTGTCTAGATTCTGGAAGT 243  
 QY 59 CysGlySerLeuValGluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78

Db 244 TTCTCTCTCTCACTGAGTTTTTCGTGAGAGAGACAGAAATCATCCAAAGCTAAGGTAGGA 303  
 Qy 79 ArgTyrLysAsnProGluGlnAlaPhePheProGluAsnLeuProSerIleVal 98  
 Db 304 AGATATGAATACCCGGAAGAGATCTTCTCTTGGAGACATTCCTCACTCGGTTTTT 363  
 Qy 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
 Db 364 CCTACGCAAAATATCCATCCGCTTTGCACCTAAGGCTCTATCTGTAAACATTAAACAA 423  
 Qy 119 SerIleTrpLysMetTyrPheLysGlnLeuLeuPheProGluAlaThrSerGlyAsp 138  
 Db 424 CAAATCTGGATATTACTTTAAAGAAATGCTTCTTTGTCAAACCTGGCGATGAT 483  
 Qy 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArg 158  
 Db 484 GGCACATATCCATCACTCTGCTAGTGTGCTGCTGCTTTTACAAAGCTTTTCGAGAAGG 543  
 Qy 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 Db 544 ATTCACTACGGTAATTTTCTAGCTGAGGTCAAAATTCAGAGATGCTCCAAAGATTACGAG 603  
 Qy 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198  
 Db 604 CTGCGATTCGGCTCAGATAGAGAGGCTTTGATGAAGCTCTTGACGTTTCAGAAAGTA 663  
 Qy 199 GluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsn 218  
 Db 664 GAAGAAATGTTTAAAGAAGAGAGTGCAGAGAAGAACGCTTTTGGACAGAGTAAAA 723  
 Qy 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236  
 Db 724 TTCACCTCGGTATGGCGATGAGAGTAAAGAGAAAGTATPAAAGTGGATCCATGCTGCC 783  
 Qy 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
 Db 784 TCTCGCATCTACGGGAATGGTATTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTT 843  
 Qy 257 LeuArgArgLeuAsp 261  
 Db 844 CTACGCGCTCGAT 858  
 RESULT 5  
 ABA91400/c  
 ID ABA91400 standard; DNA; 1006 BP.  
 XX AC ABA91400;  
 XX DT 08-APR-2002 (first entry)  
 XX DE Arabidopsis chorismate mutase-2 antisense polynucleotide.  
 XX KW Chorismate mutase-2; enzyme; plant; antisense; ss.  
 XX OS Arabidopsis thaliana.  
 XX PN WO200202798-A2.  
 XX PD 10-JAN-2002.  
 XX PF 22-JUN-2001; 2001WO-US020104.  
 XX PR 05-JUL-2000; 2000US-00610040.  
 XX PA (PARA-) PARADIGM GENETICS INC.  
 XX PI Boyes DC, Davis KR, Woessner JP, Gorchach J, Hamilton CM;  
 XX P1 Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;  
 XX P1 Kjemtrup S;  
 XX DR WPI; 2002-154754/20.  
 XX PT Identifying compounds useful for modulating and inhibiting plant growth

PT and development, by using chorismate synthase and chorismate mutase as  
 PT targets which are essential for plant growth.  
 XX Claim 35; Page 71-72; 73pp; English.

The present sequence is that of an antisense oligonucleotide that is  
 specific for Arabidopsis thaliana chorismate mutase-2 (CM) mRNA. CM  
 catalyses the conversion of chorismate to prephenate in a branch of the  
 shikimate pathway leading to phenylalanine and tyrosine biosynthesis. It  
 is essential for plant growth; inhibition of CM gene expression in plant  
 seedlings results in severe chlorosis, reduced growth and developmental  
 abnormalities. The invention provides compositions for the modulation of  
 plant growth or development using CM antisense and sense polynucleotides,  
 double-stranded RNA and ribozymes, and related expression cassettes and  
 vectors. The compositions are particularly useful for the modulation and  
 inhibition of plant growth, and for generating male sterile plants or  
 plants that produce seedless fruit

Sequence 1006 BP; 302 A; 226 C; 188 G; 290 T; 0 U; 0 Other;

Alignment Scores: Length: 1006  
 Pred. NO.: 7,06e-87 Matches: 162  
 Score: 818.00 Conservative: 31  
 Percent Similarity: 72.83% Mismatches: 68  
 Best Local Similarity: 61.13% Indels: 4  
 Query Match: 61.09% Gaps: 2  
 DB: 6

US-10-624-061-16 (1-261) x ABA91400 (1-1006)

Qy 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 Db 943 ATGCAAGAGCTCTCGAATCGGATTCGGTTCGGTTCCTCAATGTACTGAGCTTGAC 884  
 Qy 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArg 38  
 Db 883 TTAATCAGAAATCGTTGATTAGCAAGAAGACACCATCGTCTTCAGCTTGATCGAGAGA 824  
 Qy 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 Db 823 GCTAAGTTTCCACTCAATCTCTCGTTCGAGGAATCTCGTGTCTAGATTCTGGAAGT 764  
 Qy 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 Db 763 TTCTCTCTCTCACTGAGTTTTCGTGAGAGAGACAGAAATCATCCAAAGTAAAGTAGGA 704  
 Qy 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProSerIleVal 98  
 Db 703 AGATATGAATACCCGGAAGAGATCTTCTCTTGTGAGAACATCTCTCACTCGGTTTTT 644  
 Qy 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
 Db 643 CCTACGCAAAATATCCATCCGCTTTCACCCCTAAGGCTCTAICTGTATAACATTAAACAA 584  
 Qy 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138  
 Db 583 CAAATCTGGGATATTACTTTAAAGAAATGCTTCTTGTTCCTCAACCTGGCGATGAT 524  
 Qy 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158  
 Db 523 GGCACATATCCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 464  
 Qy 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 Db 463 ATTCACCTACGGTAAATTTGTAGCTGAGGTCAAATTCAGAGATGCTCCACAAAGATTACGAG 404  
 Qy 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198  
 Db 403 CTTGCGATTCGCGCTCAGATAGAGGCTTTGATGAAGCTTTGATGAGCTTTGAGAAAGTA 344  
 Qy 199 GluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsn 218  
 Db 343 GAAGAAATGTTTAAAGAAGAGAGTGCAGAGAAGAACGCTTTTGGACAGAAAGTAAAA 284

QY 219 LeuAnSerAspAsnAsnGluAnArgLysPhe-----AspProSerValAla 236  
 DB 283 TTCACCTCGCTATGGCGATGAGAGTAAGAGTAAGTAAAGTCCATTCCTGGCC 224  
 QY 237 SerSerLeuTyrLysAsnTyrValileProLeuThrLysGluValGlnValGluTyrLeu 256  
 DB 223 TCTCCATCTACGGGATGCTTATCCCTCTCACTAAGCTCTGTGAGTTGAGTACTT 164  
 QY 257 LeuArgArgLeuAsp 261  
 DB 163 CTACGTCGTCGAT 149

## RESULT 6

ABR91394  
 ID ABA91394 standard; cDNA; 1006 BP.

XX ABA91394;  
 XX 08-APR-2002 (first entry)  
 DE Arabidopsis chorismate mutase-2 cDNA.  
 XX  
 XX Chorismate mutase-2; herbicide; plant; enzyme; gene; ss.  
 XX Arabidopsis thaliana.

OS WO200202798-A2.  
 PN 10-JAN-2002.

PD 22-JUN-2001; 2001WO-US020104.  
 PF 05-JUL-2000; 2000US-00610040.  
 PR (PARA-) PARADIGM GENETICS INC.  
 PA  
 XX

PI Boyes DC, Davis KR, Woessner JP, Gorlach J, Hamilton CM;  
 PI Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Malpuri R;  
 PI Kjentrup S;  
 XX  
 DR WPI; 2002-154754/20.

XX Identifying compounds useful for modulating and inhibiting plant growth  
 PT and development, by using chorismate synthase and chorismate mutase as  
 PT targets which are essential for plant growth.

PS Claim 37; Page 64-65; 73pp; English.

XX The present sequence is that of a full-length cDNA encoding Arabidopsis  
 CC thaliana chorismate mutase-2 (CM), an enzyme that catalyses the  
 CC conversion of chorismate to prephenate in a branch of the shikimate  
 CC pathway leading to phenylalanine and tyrosine biosynthesis. CM is  
 CC essential for plant growth; inhibition of CM gene expression in plant  
 CC seedlings results in severe chlorosis, reduced growth and developmental  
 CC abnormalities. The invention provides compositions for the modulation of  
 CC plant growth or development using CM antisense (see ABA91400) and sense  
 CC polynucleotides, double-stranded RNA and ribozymes, and related  
 CC expression cassettes and vectors. The compositions are particularly  
 CC useful for the modulation and inhibition of plant growth, and for  
 CC generating male sterile plants or plants that produce seedless fruit. CM  
 CC is also a target for the identification of herbicides, and methods are  
 CC provided for using CM in the identification of herbicide candidates.  
 CC Methods are also provided for identifying compounds that stimulate the  
 CC expression or function of CM. These can be used to promote growth and  
 CC development

SQ Sequence 1006 BP; 230 A; 188 C; 226 G; 302 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 7,068-87 Length: 1006  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservative: 31

Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.03% Indels: 4  
 DB: 6 Gaps: 2

US-10-624-061-16 (1-261) x ABA91394 (1-1006)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 DB 64 ATGGCAAGAGTCTTCGATCCGATTCGGGTCTCGTGTTCATGTAAGTACTGAGTCTGAC 123  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArg 38  
 DB 124 TTAATCAGAGAATCGTTGATTAGGCAAGAGACCATCGTCTTCAGCTTGCAGATTGAGAGA 183  
 QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnLysPhe 58  
 DB 184 GCTAAGTTTCCACTCAATCTCTCTGCTTCGAGGAATCTCGTTGTCTAGATCTCGAAGT 243  
 QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 DB 244 TTCTCTTCTCTCACTGAGTTTTCGTCAGAGAGACAGAAATCATCCAAGCTAAGGTAGGA 303  
 QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleVal 98  
 DB 304 AGATATGAATACCCGGAAGAGAAATCTTCTTCTTGTGAGAACATTCCTCATCTCGTTTTT 363  
 QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
 DB 364 CTACGCACAAATATCCATCGGCTTTCACCTTTCAGGCTCTATCTGTTAACATTAAACAA 423  
 QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138  
 DB 424 CAAATCTGGGATATTACTTTAAAGATTTGCTTCTTGTGTTTGTCAAACCTGGCGATGAT 483  
 QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArg 158  
 DB 484 GGCACATATCCATCACTGCTGCTAGTATCTCGCTCTTTCAGAGCTCTTTCGAGAGG 543  
 QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 DB 544 ATTCACTACGGTAAATTTGAGTCTGAGGTCAAATTCAGAGATGCTCCACAAGATTACGAG 603  
 QY 179 ProLeuIleArgAlaLysAspLysGluLysMetLysLeuLeuThrPheThrSerVal 198  
 DB 604 CTGGCATTCGGCTCAGATAGAGGCTTTGATGAAGCTGTGACCTTTGAGAAAGTA 663  
 QY 199 GluGluThrValArgLysArgValGluLysLysAlaValPheGlyGlnGluValAsn 218  
 DB 664 GAAGAAATGTTTAAAGAGAGAGTGCAGAAAGAGAAACGTTTGGACAAGAGTAAAG 723  
 QY 219 LeuAnSerAspAsnAsnAspAsnGluAnArgLysPhe-----AspProSerValAla 236  
 DB 724 TTCAACTCTGGCTATGGCGATGAGAGTAAGAGTAAGTAAAGTCCATTCCTGCTGCC 783  
 QY 237 SerSerLeuTyrLysAsnTyrValileProLeuThrLysGluValGlnValGluTyrLeu 256  
 DB 784 TCTCCATCTACGGGATGCTTATCCCTCTCACTAAGCTCTGTGAGTTGAGTACTT 843  
 QY 257 LeuArgArgLeuAsp 261  
 DB 844 CTACGTCGTCGAT 858

## RESULT 7

ABX93053  
 ID ABX93053 standard; cDNA; 525 BP.

XX AC ABX93053;

XX 22-MAY-2003 (first entry)  
 XX Soybean chorismate mutase expressed sequence tag.

XX Soybean; plant; gene; ss; chorismate mutase; expressed sequence tag; EST;



KW aromatic amino acid; diet; biosynthesis; chorismate; anthranilate;  
KW tryptophan; prephenate; tyrosine; phenylalanine;  
KW prephenate dehydrogenase; herbicide; food crop.  
XX  
OS Glycine max.

FT Key Location/Qualifiers  
FT CDS 43..438  
FT /\*tag= a  
FT /product= "Chorismate mutase"  
FT /partial  
FT /note= "No start or stop codon shown"

PN US2002184658-A1.  
XX  
XX 05-DEC-2002.  
XX 03-DEC-1999; 99US-00454279.  
XX 04-DEC-1998; 98US-0110845P.  
XX (FALC/) FALCO S. C.  
PA (FAMC/) FAMODU O. O.  
PA (LEBJ/) LEE J.  
XX Falco SC, Famodu OO, Lee J;  
XX WPI; 2003-328651/31.  
DR P-PSDB; AB008089.  
XX  
XX New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
PT nucleic acids, useful for facilitating design and/or identifying  
PT inhibitors of those enzymes that may be used as herbicides and for  
PT producing antibodies.

XX Claim 3; Page 16; 32pp; English.  
XX  
XX The invention discloses isolated polynucleotides encoding chorismate  
CC mutase polypeptides. Aromatic amino acids must be included in the diets  
CC of animals. In the aromatic amino acid biosynthetic pathway chorismate is  
CC converted to anthranilate during tryptophan biosynthesis and is converted  
CC to prephenate, the branch point for tyrosine and phenylalanine  
CC biosynthesis. Chorismate mutase catalyses the conversion of chorismate to  
CC prephenate. Also disclosed are methods for selecting an isolated  
CC polynucleotide that affects the level of expression of a tyrosine  
CC biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
CC fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
CC evaluating compounds for their ability to inhibit the activity of a  
CC tyrosine biosynthetic enzyme. The polypeptides can be used produce  
CC antibodies. Chorismate mutase and prephenate dehydrogenase are good  
CC targets for herbicides that will not affect animals, and overexpression  
CC of these enzymes may be used to increase the content of aromatic amino  
CC acid in food crops. The polypeptides may also be used to design and/or  
CC identify inhibitors of those enzymes that may be used as herbicides. The  
CC nucleic acids may be used to create transgenic plants, as probes for the  
CC genetic and physical mapping of the genes and as markers for traits  
CC linked to those genes. The sequence presented is a soybean chorismate  
CC mutase expressed sequence tag (EST)  
XX  
XX Sequence 525 BP; 160 A; 103 C; 111 G; 148 T; 0 U; 3 Other;

Alignment Scores:  
Pred. No.: 2,72e-81 Length: 525  
Score: 767.50 Matches: 162  
Percent Similarity: 95.29% Conservative: 0  
Best Local Similarity: 95.29% Mismatches: 7  
Query Match: 57.32% Indels: 4  
DB: 7 Gaps: 0

US-10-624-061-16 (1-261) x ABX93053 (1-525)  
Qy 1 MetAlaLysAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
|||||

Db 16 ATGCCCAAGCAGCAGACACAAAGTCTGATCTGGGAATGTGTACACGCTAGCTTTCTGTG 75  
Qy 21 ArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40  
Db 76 AGAGAGGATTGGTTAGGCAAGAGGATACCACTTTAIGGTTCTCATTTAGAGAGGCAAG 135  
Qy 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
Db 136 TTCCTAGCAATTTCTCACACCTATGATGAAGAAGTATGCTCAATCCAGGGTTTTTGTGGC 195  
Qy 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
Db 196 TCATTGGTGAATTTGTTTGAATAAGAGGAGGATCAAGCTTAAGCTGAAGTGAAGATAC 255  
Qy 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100  
Db 256 AAAAAACCTGAGAAAAAGCCITTTCCAGAAAAATTTACCACCATCAATTTGCGCATCT 315  
Qy 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
Db 316 TACTCCTTCAACACAGTTTTTGTCATCTCTGGTGGCTTCAATTAACATAACAAGTCACTCT 375  
Qy 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140  
Db 376 GGGAAATGATTTCAGAGAGTTACTTCCATTGCTTCTGCTACITCGGTGATGATGNAAC 435  
Qy 141 TyrAlaGln-ThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgIleHi 160  
Db 436 TATGCGCAAACTGCAGCTAATGACCT-TCAATTTATTCAGTTC-ATCTCTAGAAGGAT-CA 492  
Qy 160 STYRGlyVysPheValAlaGluValLys 169  
Db 493 CTATGGAAGTTTGTAGCTGANGNGAAA 520  
RESULT 8  
ABZ12932  
ID ABZ12932 standard; DNA; 1005 BP.  
XX  
AC ABZ12932;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 737.  
XX  
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200216655-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 24-AUG-2001; 2001WO-US026685.  
XX  
PR 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.  
XX  
XX (SRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Harper JF, Kreps J, Wang X, Zhu T;  
XX  
DR WPI; 2002-304127/34.  
XX  
XX Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.  
XX  
PS Claim 144; SEQ ID NO 737; 577pp + Sequence Listing; English.  
XX  
CC The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an

CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office  
 XX  
 XX Sequence 1005 BP; 287 A; 205 C; 235 G; 278 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1.17e-72 Length: 1005  
 Score: 698.00 Matches: 131  
 Percent Similarity: 68.08% Conservative: 46  
 Best Local Similarity: 50.38% Mismatches: 75  
 Query Match: 52.13% Indels: 8  
 DB: 6 Gaps: 1

US-10-624-061-16 (1-261) x ABZ12932 (1-1005)

QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29  
 DB 223 GATGAGAGTGAGGTTGACTCTTGAAGGTATTGAACTCTTTGATCCGTCAAGAGGAC 282  
 QY 30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49  
 DB 283 AGCATTATATTGGCTATTGGAGAGGCCAAGTACTGTTACAATGCTGATCTATGAT 342  
 QY 50 GluLysTyrAlaGlnIleGlyPheCysGlySerLeuValGluPheValValLysAsn 69  
 DB 343 CCTACTGCTTTGATGATGATGTTTCAATGCTTCTTGGTGGTACATGTTGTTAAAGGC 402  
 QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89  
 DB 403 ACTGAGAGCTTCAGCTTAAGGTTGGTAGTTTAAAGTCTGATGAACATCTTCTTC 462  
 QY 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
 DB 463 CCTGATGATCACCAGAGCCCTATGTTGCTCTCTTCAGTACCACAAAGGTGTTCATTT 522  
 QY 110 GlyAlaAlaSerIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129  
 DB 523 GCTGCTGATTCGATAAACAACAAGAGATGGAACATGACTTACAGACCTTGT 582  
 QY 130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
 DB 583 CCAAGACTTGTGAAGAAGGCGATGATGTAACCTACCGCTCAACAGCTGTCTGTGACGCT 642  
 QY 150 SerLeuLeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLys 169  
 DB 643 ATTCGCTTCTGCTCTCTCAAGAGATTCATTCAGCTTAATTTGTTGAGAGCTAA 702  
 QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGlyLeu 189  
 DB 703 TTTCAAGCTCACCGAAGCATACGAGTCCGCCATCAAGCACAGATAGGATGACCTG 762  
 QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209  
 DB 763 ATGATATGCTGATTCCTCCGACTGTGGAAGATCGGATAAAGAGAGAGTTGATGAA 822  
 QY 210 AlaValValPheGlyGlnGluValAlaAsnLysSerAspAsnAspGlnGluAsnArg 229  
 DB 823 ACCCGAACATACGGGCAAGAGTGAAGTTGGATGGAGGAGAGAGAGAGAGAGAA 882  
 QY 230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241  
 DB 883 GAAGGGAATGATCTCATGTTTCAAAATCAGTCCGATCTTATGTTGGTACTTATGGA 942  
 QY 242 AsnTrpValIleProLeuThrLysGluValGlnValGluLysLeuValArgLeuAsp 261  
 DB 943 GATTGGATCATGCTTTTAAACAAAGAGGTTCAAGTGGAGTACTTGTCTCAGAGACTGGAC 1002

## RESULT 9

ABA91393  
 ID ABA91393 standard; cDNA; 1207 BP.

XX ABA91393;  
 AC ABA91393;  
 DT 08-APR-2002 (first entry)  
 XX Arabidopsis chorismate mutase-1 cDNA.  
 XX Chorismate mutase-1; herbicide; plant; enzyme; gene; ss.

XX Arabidopsis thaliana.

XX WO200202798-A2.

XX 10-JAN-2002.

XX 22-JUN-2001; 2001WO-US020104.

XX 05-JUL-2000; 2000US-00610040.

XX (PARA-) PARADIGM GENETICS INC.

XX Boyes DC, Davis KR, Woessner JP, Gorlach J, Hamilton CM;  
 PI Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulipuri R;  
 PI Kjemtrup S;  
 XX WPI; 2002-154754/20.

XX Claim 37; Page 64; 73pp; English.

XX Identifying compounds useful for modulating and inhibiting plant growth  
 PT and development, by using chorismate synthase and chorismate mutase as  
 PT targets which are essential for plant growth.

XX PS 2002-154754/20.

XX The present sequence is that of a full-length cDNA encoding Arabidopsis  
 CC thaliana chorismate mutase-1 (CM), an enzyme that catalyses the  
 CC conversion of chorismate to prephenate in a branch of the shikimate  
 CC pathway leading to phenylalanine and tyrosine biosynthesis. CM is  
 CC essential for plant growth; inhibition of CM gene expression in plant  
 CC seedlings results in severe chlorosis, reduced growth and developmental  
 CC abnormalities. The invention provides compositions for the modulation of  
 CC plant growth or development using CM antisense (see ABA91399) and sense  
 CC polynucleotides, double-stranded RNA and ribozymes, and related  
 CC expression cassettes and vectors. The compositions are particularly  
 CC useful for the modulation and inhibition of plant growth, and for  
 CC generating male sterile plants or plants that produce seedless fruit. CM  
 CC is also a target for the identification of herbicides, and methods are  
 CC provided for using CM in the identification of herbicide candidates.  
 CC Methods are also provided for identifying compounds that stimulate the  
 CC expression or function of CM. These can be used to promote growth and  
 CC development

XX Sequence 1207 BP; 343 A; 232 C; 278 G; 354 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2.61e-72 Length: 1207  
 Score: 696.00 Matches: 131  
 Percent Similarity: 68.08% Conservative: 46  
 Best Local Similarity: 50.38% Mismatches: 75  
 Query Match: 51.98% Indels: 8  
 DB: 6 Gaps: 1

US-10-624-061-16 (1-261) x ABA91393 (1-1207)

QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29  
 DB 235 GATGAGAGTGAGGTTGACTCTTGAAGGTATTGAACTCTTTGATCCGTCAAGAGGAC 294  
 QY 30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49  
 DB 943 GATTGGATCATGCTTTTAAACAAAGAGGTTCAAGTGGAGTACTTGTCTCAGAGACTGGAC 1002

Db 295 AGCATTATTTGGCTATTGGAGAGAGCCAAAGTACTGTTACAACTGCTGATCTTATGAT 354  
 QY 50 GlutylsTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValysAsn 69  
 Db 355 CCTACTGCTTTGACATGGATGGTTCAATGGTTCTTTGGTTGAGTACATGGTTAAAGGC 414  
 QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89  
 Db 415 ACTGAGAAGCTTCAGCTAAGGTTGGTAGGTTTAAAGAGTCTGATGAACATCTTCTTC 474  
 QY 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
 Db 475 CTTGATGCTTACAGAGCTTATGCTCTCTCAAGAGAAATCCATTACGGTAAATTTGTCGAGAAGCTAAA 714  
 QY 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrPheLysGluLeuLeu 129  
 Db 535 GCTGCTGATTTCGATAAACAATAACAAGAAGATATGGAACATGATACITTCAGAGACCTTGTT 594  
 QY 130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
 Db 595 CCAAGACTTGTGAAGAAGGGGATGATGGTAACCTACCGCTCAACAGCTGTCTGTGACGCT 654  
 QY 150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys 169  
 Db 655 ATCTGCTTCACTGCTCTCAAGAGAAATCCATTACGGTAAATTTGTCGAGAAGCTAAA 714  
 QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189  
 Db 715 TTTCAAGCCTCACCAGCAAGCATACAGTCCGCCATCAAGCAAGCAAGATGAAGTCAAGCTG 774  
 QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209  
 Db 775 ATGGATATGCTGACATTCCTCCGACTGTGGAAGATCGCATAAAGAGAGAGTGTGATGAAA 834  
 QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAsnAspAsnGluAsnArg 229  
 Db 835 ACCCGAATACAGGCAAGCAAGTGAAGTGGAGTGGAGGAGAAAGAAAGAAAGAAAGAA 894  
 QY 230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241  
 Db 895 GAAGGGAATGAATCTCATGTTTACAAATTCAGTCCGATCTTACTGTTGATATATGCA 954  
 QY 242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261  
 Db 955 GATTGGATCATGCTTTTAAACAAAGAGGTTCAAGTGGAGTACTTGTCTCAAGAGACTGGAC 1014  
 RESULT 10  
 ID ABA91399/c  
 XX ABA91399;  
 AC ABA91399;  
 XX 08-APR-2002 (first entry)  
 XX Arabidopsis chorismate mutase-1 antisense polynucleotide.  
 DE Chorismate mutase-1; enzyme; plant; antisense; ss.  
 KW Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 FN WO200202798-A2.  
 PD 10-JAN-2002.  
 PF 22-JUN-2001; 2001WO-US020104.  
 PR 05-JUL-2000; 2000US-00610040.  
 XX (PARA-) PARADIGM GENETICS INC.  
 XX Boyes DC, Davis KR, Woessner JP, Gortlach J, Hamilton CM;  
 PI Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;  
 PI Kjemtrup S;

XX

DR

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

WPI; 2002-154754/20.

Identifying compounds useful for modulating and inhibiting plant growth and development, by using chorismate synthase and chorismate mutase as targets which are essential for plant growth.

Claim 35; Page 71; 73pp; English.

The present sequence is that of an antisense oligonucleotide that is specific for Arabidopsis thaliana chorismate-mutase-1 (CM) mRNA. CM catalyzes the conversion of chorismate to prephenate in a branch of the shikimate pathway leading to phenylalanine and tyrosine biosynthesis. It is essential for plant growth; inhibition of CM gene expression in plant seedlings results in severe chlorosis, reduced growth and developmental abnormalities. The invention provides compositions for the modulation of plant growth or development using CM antisense and sense polynucleotides, double-stranded RNA and ribozymes, and related expression cassettes and vectors. The compositions are particularly useful for the modulation and inhibition of plant growth, and for generating male sterile plants or plants that produce seedless fruit

Sequence 1207 BP; 354 A; 278 C; 232 G; 343 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,61e-72 Length: 1207  
 Score: 696.00 Matches: 131  
 Percent Similarity: 68.08% Conservative: 46  
 Best Local Similarity: 50.38% Mismatches: 75  
 Query Match: 51.98% Indels: 8  
 DB: 6 Gaps: 1

US-10-624-061-16 (1-261) x ABA91399 (1-1207)

QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29

Db 973 GATGAGAGTGCAGAGTTTGACTCTTGAAGGTATTATGAAGTCTTTTGATCCGTCAAGAGGAC 914

QY 30 ThrIleIleTyrGlyLeuIleGluAtgAlaLysPheProSerAsnSerHisThrTyrAsp 49

Db 913 AGCATATATTTGGCTATTGGAGAGGCAAGTACTGTACATGCTGATCTATGAT 854

QY 50 GlutylsTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69

Db 853 CCTACTGCTTTTGCATGATGATGTTTCAATGGTCTTCTTGGTTGAGTACATGTTAAGGC 794

QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89

Db 793 ACTGAGAGCTTCAGCTAAGTGTGGTAGGTTTAAAGTCTCTGATGAAACATCTTCTTC 734

QY 90 ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109

Db 733 CCTGATGATCTACGAGAGCTATGTTGCTCTCTTCACTAGTACCAAGGTTTGCATTTT 674

QY 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129

Db 673 GCTGCTGATTCGATTAACATAAACAAGAGATATGGAACATGATCTTACAGACCTTGT 614

QY 130 ProLeuLeuAlaThrSerGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149

Db 613 CCAAGACTTGTGAAGAAAGGCGATGATGTTAACTACGCTCAACAGCTCTCTGTGACGCT 554

QY 150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys 169

Db 553 ATCTGCTTCTGCTCTCTCAAGAGAAATCCATTACGCTAAATTTGTTCCAGAGCTAAA 494

QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189

Db 493 TTTCAAGCCTTCACCCGAAAGCATACGAGTCCGCCATCAAGAGCACAAGATAAGGATCGACTG 434

QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209

Db 433 ATGGATATGCTGACATTCGCGACTGTGGAGATGCGATTAAGAGAGAGTGTGATGATA 374



QY	260	Leuasp 261	PR	18-JUN-1999;	99US-0139455P.
			PR	18-JUN-1999;	99US-0139456P.
DB	1018	TTGGAT 1023	PR	18-JUN-1999;	99US-0139457P.
			PR	18-JUN-1999;	99US-0139458P.
RESULT 12			PR	18-JUN-1999;	99US-0139459P.
AAC33324			PR	18-JUN-1999;	99US-0139460P.
ID AAC33324		standard; DNA; 1274 BP.	PR	18-JUN-1999;	99US-0139461P.
XX			PR	18-JUN-1999;	99US-0139462P.
AC			PR	18-JUN-1999;	99US-0139463P.
AAC33324;			PR	18-JUN-1999;	99US-0139750P.
DT	17-OCT-2000	(first entry)	PR	18-JUN-1999;	99US-0139763P.
XX			PR	21-JUN-1999;	99US-0139817P.
DE		Arabidopsis thaliana DNA fragment SEQ ID NO: 2614.	PR	22-JUN-1999;	99US-0139899P.
XX			PR	23-JUN-1999;	99US-0140353P.
XX			PR	23-JUN-1999;	99US-0140354P.
KW		Hybridisation assay; genetic mapping; gene expression control;	PR	24-JUN-1999;	99US-0140895P.
KW		protein identification; signal transduction pathway; metabolic pathway;	PR	28-JUN-1999;	99US-0140823P.
KW		promoter; termination sequence; ss.	PR	29-JUN-1999;	99US-0140991P.
XX			PR	30-JUN-1999;	99US-0141287P.
OS		Arabidopsis thaliana.	PR	01-JUL-1999;	99US-0141842P.
XX			PR	01-JUL-1999;	99US-0142154P.
PN		EP1033405-A2.	PR	02-JUL-1999;	99US-0142055P.
XX			PR	06-JUL-1999;	99US-0142390P.
PD		06-SEP-2000.	PR	08-JUL-1999;	99US-0142803P.
XX			PR	09-JUL-1999;	99US-0142920P.
PF		25-FEB-2000; 2000EP-00301439.	PR	12-JUL-1999;	99US-0142977P.
XX			PR	13-JUL-1999;	99US-0143542P.
PR		25-FEB-1999;	PR	14-JUL-1999;	99US-0143624P.
PR		99US-0121825P.	PR	15-JUL-1999;	99US-0144005P.
PR		99US-0123180P.	PR	16-JUL-1999;	99US-0144085P.
PR		09-MAR-1999;	PR	16-JUL-1999;	99US-0144086P.
PR		99US-0123548P.	PR	19-JUL-1999;	99US-0144325P.
PR		23-MAR-1999;	PR	19-JUL-1999;	99US-0144331P.
PR		99US-0125788P.	PR	19-JUL-1999;	99US-0144332P.
PR		25-MAR-1999;	PR	19-JUL-1999;	99US-0144333P.
PR		99US-0126264P.	PR	19-JUL-1999;	99US-0144334P.
PR		29-MAR-1999;	PR	19-JUL-1999;	99US-0144335P.
PR		99US-0126785P.	PR	20-JUL-1999;	99US-0144352P.
PR		01-APR-1999;	PR	20-JUL-1999;	99US-0144632P.
PR		99US-0127462P.	PR	20-JUL-1999;	99US-0144884P.
PR		06-APR-1999;	PR	21-JUL-1999;	99US-0145086P.
PR		99US-0128714P.	PR	21-JUL-1999;	99US-0145088P.
PR		16-APR-1999;	PR	21-JUL-1999;	99US-0145089P.
PR		99US-0130077P.	PR	22-JUL-1999;	99US-0145192P.
PR		21-APR-1999;	PR	22-JUL-1999;	99US-0145145P.
PR		99US-0129845P.	PR	22-JUL-1999;	99US-0145218P.
PR		19-APR-1999;	PR	23-JUL-1999;	99US-0145224P.
PR		99US-0130449P.	PR	23-JUL-1999;	99US-0145226P.
PR		23-APR-1999;	PR	26-JUL-1999;	99US-0145276P.
PR		99US-0130891P.	PR	27-JUL-1999;	99US-0145913P.
PR		28-APR-1999;	PR	27-JUL-1999;	99US-0145918P.
PR		99US-0131449P.	PR	27-JUL-1999;	99US-0145919P.
PR		30-APR-1999;	PR	28-JUL-1999;	99US-0145951P.
PR		99US-0132048P.	PR	02-AUG-1999;	99US-0146386P.
PR		04-MAY-1999;	PR	02-AUG-1999;	99US-0146388P.
PR		99US-0132484P.	PR	02-AUG-1999;	99US-0146389P.
PR		05-MAY-1999;	PR	03-AUG-1999;	99US-0147038P.
PR		99US-0132485P.	PR	04-AUG-1999;	99US-0147204P.
PR		06-MAY-1999;	PR	04-AUG-1999;	99US-0147302P.
PR		99US-0132486P.	PR	05-AUG-1999;	99US-0147192P.
PR		06-MAY-1999;	PR	05-AUG-1999;	99US-0147260P.
PR		99US-0132487P.	PR	06-AUG-1999;	99US-0147303P.
PR		07-MAY-1999;	PR	06-AUG-1999;	99US-0147416P.
PR		99US-0132863P.	PR	09-AUG-1999;	99US-0147935P.
PR		11-MAY-1999;	PR	09-AUG-1999;	99US-0148171P.
PR		99US-0134218P.	PR	10-AUG-1999;	99US-0148319P.
PR		14-MAY-1999;	PR	11-AUG-1999;	99US-0148341P.
PR		99US-0134219P.	PR	12-AUG-1999;	99US-0148565P.
PR		14-MAY-1999;	PR	13-AUG-1999;	99US-0148684P.
PR		99US-0134370P.	PR	13-AUG-1999;	99US-0149368P.
PR		14-MAY-1999;	PR	16-AUG-1999;	
PR		99US-0134768P.			
PR		18-MAY-1999;			
PR		99US-0134941P.			
PR		19-MAY-1999;			
PR		99US-0135124P.			
PR		20-MAY-1999;			
PR		99US-0135353P.			
PR		21-MAY-1999;			
PR		99US-0135629P.			
PR		24-MAY-1999;			
PR		99US-0136021P.			
PR		25-MAY-1999;			
PR		99US-0136392P.			
PR		27-MAY-1999;			
PR		99US-0136782P.			
PR		28-MAY-1999;			
PR		99US-0137222P.			
PR		01-JUN-1999;			
PR		99US-0137528P.			
PR		03-JUN-1999;			
PR		99US-0137502P.			
PR		04-JUN-1999;			
PR		99US-0137724P.			
PR		07-JUN-1999;			
PR		99US-0138094P.			
PR		08-JUN-1999;			
PR		99US-0138540P.			
PR		10-JUN-1999;			
PR		99US-0138847P.			
PR		10-JUN-1999;			
PR		99US-0139119P.			
PR		14-JUN-1999;			
PR		99US-0139452P.			
PR		16-JUN-1999;			
PR		99US-0139453P.			
PR		17-JUN-1999;			
PR		99US-0139492P.			
PR		18-JUN-1999;			
PR		99US-0139454P.			



US2002184658-A1.

05-DEC-2002.

03-DEC-1999; 99US-00454279.

04-DEC-1998; 98US-0110845P.

(FALC/) FALCO S. C.  
(FAMO/) FAMODU O. O.  
(LEEJ/) LEE J.

Falco SC, Famodu OO, Lee J;  
WPI; 2003-328651/31.  
P-PSDB; ABU08092.

New tyrosine biosynthetic enzyme, chorismate dismutase and  
PT nucleic acids, useful for facilitating design and/or identifying  
PT inhibitors of those enzymes that may be used as herbicides and for  
PT producing antibodies.

Claim 3; Page 19; 32pp; English.

The invention discloses isolated polynucleotides encoding chorismate  
mutase polypeptides. Aromatic amino acids must be included in the diets  
of animals. In the aromatic amino acid biosynthetic pathway chorismate is  
converted to anthranilate during tryptophan biosynthesis and is converted  
to prephenate, the branch point for tyrosine and phenylalanine  
biosynthesis. Chorismate mutase catalyses the conversion of chorismate to  
prephenate. Also disclosed are methods for selecting an isolated  
polynucleotide that affects the level of expression of a tyrosine  
biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
evaluating compounds for their ability to inhibit the activity of a  
tyrosine biosynthetic enzyme. The polypeptides can be used produce  
antibodies. Chorismate mutase and prephenate dehydrogenase are good  
targets for herbicides that will not affect animals, and overexpression  
of these enzymes may be used to increase the content of aromatic amino  
acid in food crops. The polypeptides may also be used to design and/or  
identify inhibitors of those enzymes that may be used as herbicides. The  
nucleic acids may be used to create transgenic plants, as probes for the  
genetic and physical mapping of the genes and as markers for traits  
linked to those genes. The sequence presented is a corn chorismate mutase  
cDNA

Sequence 1223 BP; 341 A; 284 C; 312 G; 286 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,79e-69	Length:	1223
Score:	670.50	Matches:	130
Percent Similarity:	70.36%	Conservative:	48
Best Local Similarity:	51.38%	Mismatches:	74
Query Match:	50.07%	Indels:	1
DB:	7	Gaps:	1

US-10-624-061-16 (1-261) x ABX93056 (1-1223)

10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29

237 GATCGAAGTGACATATTGCATTCGATAGCATCAGACAAAGTTTGTAGACTAGAGAC 296

30 ThrIleuTyrGlyLeuIleuIduArgAlaLysPheProSerAnSerHisThrTyrAsp 49

297 AGCATCATATTGGCCCTTTTGAGAGAGACACAGTTTGTTCACAATGCTGATACATATGAT 356

50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValLysAsn 69

357 AGCAATGCTTCCACATGATGCTTTTGGAGAGACTTTGGTGTATATATAGTTAGAGAA 416

70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrIlysAsnProGluGluAsnAlaPhePhe 89

417 ACTGAAGAAGTCATCGACAGGTCGGGAGATACAAAGCCAGAGCCAGATGAGACCCCTTTCTTT 476

CC The present sequence is that of a full-length cDNA encoding Arabidopsis  
 CC thaliana chorismate mutase-3 (CM), an enzyme that catalyses the  
 CC conversion of chorismate to prephenate in a branch of the shikimate  
 CC pathway leading to phenylalanine and tyrosine biosynthesis. CM is  
 CC essential for plant growth; inhibition of CM gene expression in plant  
 CC seedlings results in severe chlorosis, reduced growth and developmental  
 CC abnormalities. The invention provides compositions for the modulation of  
 CC plant growth or development using CM antisense (see ABA91401) and sense  
 CC polynucleotides, double-stranded RNA and ribozymes, and related  
 CC expression cassettes and vectors. The compositions are particularly  
 CC useful for the modulation and inhibition of plant growth, and for  
 CC generating male sterile plants or plants that produce seedless fruit. CM  
 CC is also a target for the identification of herbicides, and methods are  
 CC provided for using CM in the identification of herbicide candidates.  
 CC Methods are also provided for identifying compounds that stimulate the  
 CC expression or function of CM. These can be used to promote growth and  
 CC development  
 XX  
 SQ Sequence 1217 BP; 349 A; 273 C; 235 G; 360 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,18e-67 Length: 1217  
 Score: 654.50 Matches: 128  
 Percent Similarity: 67.98% Conservative: 44  
 Best Local Similarity: 50.59% Mismatches: 80  
 Query Match: 48.88% Indels: 1  
 DB: 6 Gaps: 1

US-10-624-061-16 (1-261) x ABA91395 (1-1217)

QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGluAsp 29  
 DB 332 GATGAGTGTGAGTATTGAACTTGAACCTGAAGCACTAGACACTCTTTGATTGTCGTCAGAGGAC 391  
 QY 30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49  
 DB 392 AGTATATCTTTAACTCTTCTTGAACGAGCTCAGTACCTCAACGCTGATACCTATTATGAC 451  
 QY 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69  
 DB 452 GAGGATGCTTTACTATGGAAGGTTTCAGGAATCTTTAGTTAGTTATGTCAGAGAA 511  
 QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89  
 DB 512 ACTGAAAGCTTCACGCAAGGTGACAGGTACAGAGTCTGCTGATGACATCCCTTTTC 571  
 QY 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
 DB 572 CCACATGCTTGCCTGACCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 631  
 QY 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTrpLysMetTyrPheLysGluLeuLeu 129  
 DB 632 TCGCGCGAATCGATAAATCAATCAACAAGAGGTGGAATATGTATTTCAAAACACCTTCTC 691  
 QY 130 ProLeuLeuAlaThrSerClyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
 DB 692 CCCAGACTGGTCAACCCAGGGGATGACGGAATTTGTTGTTGTTGTTGTTGTTGTTGTTG 751  
 QY 150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys 169  
 DB 752 ATGTGTTTGCAGATACCTTTCAAGAGAGATTCACCTGCGTAATTTGTTGTCGCGCAAG 811  
 QY 170 PheArgAspAlaProGluAspTyrGluProIleAlaLysAspLysGluGlyLeu 189  
 DB 812 TTTCTGTGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 871  
 QY 190 MetLysLeuLeuThrPheThrSerValGluThrValArgLysArgValGluLysLys 209  
 DB 872 ATGCAACTCTTAACGTACGACGAGCTTCAAGAGTAGTCAAGAGAGAGAGAGAGAGAGAG 931  
 QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAsnAsp---AsnGluAsn 228  
 DB 932 GCCAGAAATTTTGGTCAAGACATACGATTAACGACCCAGCAACCTGAAGCTGATCTCTTC 991

QY 229 ArgLysPheAspProSerValAlaSerSerLeuTyrLysAsnTrpValIleProLeuThr 248  
 DB 992 TACAAAATACACACTAGCTTAGTTGCAAACTCTATGGAGAAAGGATCATGCCCTCACA 1051  
 QY 249 LysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261  
 DB 1052 AAGGAAGTCAAAATGAGTACTTGTCTTAGAAGACTGGAT 1090

RESULT 15  
 ABA91401/c  
 ID ABA91401 standard; DNA; 1217 BP.  
 XX ABA91401;  
 AC ABA91401;  
 XX  
 DT 08-APR-2002 (first entry)  
 XX  
 DE Arabidopsis chorismate mutase-3 antisense polynucleotide.  
 XX  
 KW Chorismate mutase-3; enzyme; plant; antisense; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN W0200202798-A2.  
 XX  
 PD 10-JAN-2002.  
 XX  
 PF 22-JUN-2001; 2001WO-US020104.  
 XX  
 PR 05-JUL-2000; 2000US-00610040.  
 XX  
 PA (PARA-) PARADIGM GENETICS INC.  
 XX  
 PI Boyes DC, Davis KB, Woessner JP, Gorlach J, Hamilton CM;  
 PI Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;  
 PI Kjemtrup S;  
 XX  
 DR WPI; 2002-154754/20.

XX Identifying compounds useful for modulating and inhibiting plant growth  
 PT and development, by using chorismate synthase and chorismate mutase as  
 PT targets which are essential for plant growth.  
 XX  
 PS Claim 35; Page 72; 73pp; English.  
 XX  
 CC The present sequence is that of an antisense oligonucleotide that is  
 CC specific for Arabidopsis thaliana chorismate-mutase-2 (CM) mRNA. CM  
 CC catalyses the conversion of chorismate to prephenate in a branch of the  
 CC shikimate pathway leading to phenylalanine and tyrosine biosynthesis. It  
 CC is essential for plant growth; inhibition of CM gene expression in plant  
 CC seedlings results in severe chlorosis, reduced growth and developmental  
 CC abnormalities. The invention provides compositions for the modulation of  
 CC plant growth or development using CM antisense and sense polynucleotides,  
 CC double-stranded RNA and ribozymes, and related expression cassettes and  
 CC vectors. The compositions are particularly useful for the modulation and  
 CC inhibition of plant growth, and for generating male sterile plants or  
 CC plants that produce seedless fruit  
 XX

SQ Sequence 1217 BP; 360 A; 235 C; 273 G; 349 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,18e-67 Length: 1217  
 Score: 654.50 Matches: 128  
 Percent Similarity: 67.98% Conservative: 44  
 Best Local Similarity: 50.59% Mismatches: 80  
 Query Match: 48.88% Indels: 1  
 DB: 6 Gaps: 1

US-10-624-061-16 (1-261) x ABA91401 (1-1217)

QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGluAsp 29  
 DB 886 GATGAGTGTGAGTATTGAACTTGAAGCACTTAGACACTCTTTGATTGTCGTCAGAGGAC 827







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 25, 2004, 02:59:07 ; Search time 90 seconds

(without alignments)  
1609.358 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKAAEQSPDSGNVYTLASV.....NWVPLTRKVEQVEVILLRLD 261

Scoring table:

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10624061/runat.24092004.092407.16935/app.query.fasta.1.455  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFCIL=0  
-LOOPEXT=0 -UNITS=BITS -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10624061 -CGN 1.1.69 -runat.24092004.092407.16935 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.\*  
1: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/prodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/prodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/prodata/2/ina/6C COMB.seq.\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description	
1	1339	100.0	1020	4	US-09-454-279-15	Sequence 15, Appl	
2	818	61.1	1006	4	US-09-610-040-3	Sequence 3, Appl	
C	3	818	61.1	1006	4	US-09-610-040-9	Sequence 9, Appl
	4	767.5	57.3	525	4	US-09-454-279-5	Sequence 5, Appl
5	696	52.0	1207	4	US-09-610-040-2	Sequence 2, Appl	
	6	696	52.0	1207	4	US-09-610-040-8	Sequence 8, Appl
7	685.5	51.2	1231	4	US-09-454-279-13	Sequence 13, Appl	
	8	670.5	50.1	1223	4	US-09-454-279-11	Sequence 11, Appl
9	654.5	48.9	1217	4	US-09-610-040-4	Sequence 4, Appl	
	10	654.5	48.9	1217	4	US-09-610-040-10	Sequence 10, Appl
C	11	602	45.0	780	4	US-09-454-279-17	Sequence 17, Appl
	12	587	43.8	5176	4	US-09-610-040-6	Sequence 6, Appl

	13	549.5	41.0	5099	4	US-09-610-040-5	Sequence 5, Appli
	14	522	39.0	579	4	US-09-454-279-7	Sequence 7, Appli
	15	144.5	10.8	601	4	US-09-454-279-3	Sequence 3, Appli
	16	130	9.7	552	4	US-09-454-279-1	Sequence 1, Appli
C	17	105	7.8	19250	4	US-08-961-527-35	Sequence 35, Appli
	18	103	7.7	1830121	4	US-09-557-884-1	Sequence 1, Appli
	19	103	7.7	1830121	4	US-09-643-990A-1	Sequence 1, Appli
	20	96	7.2	1395	4	US-09-134-001C-1864	Sequence 1864, App
	21	96	7.2	2262	4	US-09-543-681A-850	Sequence 850, App
	22	92	6.9	3089	4	US-09-016-434-1330	Sequence 1330, App
	23	91.5	6.8	2394	3	US-09-319-989-7	Sequence 7, Appli
	24	90.5	6.8	2065	3	US-09-319-989-5	Sequence 5, Appli
	25	90.5	6.8	3081	3	US-09-319-989-9	Sequence 9, Appli
	26	90.5	6.8	3360	3	US-09-319-989-8	Sequence 8, Appli
	27	89.5	6.7	1773	4	US-09-107-532A-1028	Sequence 1028, App
	28	89	6.6	2427	4	US-09-543-681A-823	Sequence 823, App
	29	86.5	6.5	2793	4	US-09-134-001C-2477	Sequence 2477, App
	30	86.5	6.5	2976	4	US-09-386-962C-7	Sequence 7, Appli
	31	86	6.4	7953	4	US-08-956-171E-152	Sequence 152, App
	32	83.5	6.2	3414	1	US-07-973-320-1	Sequence 1, Appli
	33	83	6.2	2310	4	US-09-540-236-1190	Sequence 1190, App
	34	83	6.2	2709	4	US-09-107-532A-183	Sequence 183, App
	35	83	6.2	2749	1	US-08-344-695-1	Sequence 1, Appli
	36	83	6.2	3213	2	US-08-633-770A-4	Sequence 4, Appli
	37	83	6.2	3213	4	US-09-280-197-8	Sequence 8, Appli
	38	83	6.2	7425	4	US-09-453-702B-212	Sequence 212, App
C	39	83	6.2	119211	4	US-09-596-002-40	Sequence 40, Appli
	40	82	6.1	3414	1	US-07-973-320-3	Sequence 3, Appli
	41	81.5	6.1	2286	4	US-09-112-580-1	Sequence 1, Appli
	42	81.5	6.1	580073	4	US-08-545-528D-1	Sequence 1, Appli
	43	81	6.0	1205	4	US-09-976-594-457	Sequence 457, App
	44	81	6.0	2559	4	US-09-071-035-325	Sequence 325, App
	45	81	6.0	2559	4	US-09-071-035-329	Sequence 329, App

ALIGNMENTS

RESULT 1  
US-09-454-279-15  
; Sequence 15, Application US/09454279  
; Patent No. 6627798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Lee Jian-Ming  
; APPLICANT: Famodu, Omolayo O.  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BBI299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 15  
; LENGTH: 1020  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-454-279-15

Alignment Scores:	5.68e-163	Length:	1020
Pred. No.:	1339.00	Matches:	261
Score:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-624-061-16 (1-261) x US-09-454-279-15 (1-1020)

QY 1 MetalAllysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20

DB 23 ATGGCCAAAGCAGACAAAGTCTGATTCCTGGGAATGTGTACACGCTAGCTTCGTG 82

```

QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLys 40
Db 83 AGAGAGATTGGTTAGCAAGAGATACCATTTATGGTCTCATTTAGAGAGCCCAAG 142
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60
Db 143 TTCCTAGCAATCTCACACCTATGATGAAGAGTATGCTCAATCCAGGGTTTTGTGGC 202
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80
Db 203 TCATTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAGCTAAGCTGGAAGATAC 262
QY 81 LysAsnProGluGluAlaPhePheProGluAsnLeuProProSerIleValProSer 100
Db 263 AAAACCCCTGAAGAAACGCTTCTCCAGAAATTTACCAACCATCAATTGTGCCATCT 322
QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120
Db 323 TACTCTCTCAACACAGTTTGTGATCCTGCTGCTGCTTCAATTAACATTAACAAGTCCATC 382
QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsn 140
Db 383 TGGAAATATGATTTCAAGAGATTACTTCATCTGCTTGCTACTTCCGGTGATGATGGCAAC 442
QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgIleHis 160
Db 443 TATGGCAAACTGCAGCTAATGACCTTTCAATTATTCAGTCCATCTCTAGAAGGATTAC 502
QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180
Db 503 TATGGAATGTTTGTAGCTGAGGTGAATTCAGGATGCTCTCAAGACTACGAGCCCTTA 562
QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200
Db 563 ATTCGAGCTAAGATAAAGAGAGATTGATGAAATTTGTTGACATTTACAAGCGTTGAAGAG 622
QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220
Db 623 ACGGTGAGGAAGAGAGATTGAAAAGAGGCTGTGGTGTGTTGGCGAGGAGTGAATCTTAAC 682
QY 221 SerAspAspAsnAspGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240
Db 683 AGTGATGACATGACAAATGAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTTGATC 742
QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeu 260
Db 743 AAAAAATGGGTGATACCTCTCACCAGGAGGTTCCAGGTGAGTACCTCTTGGCGCGCTA 802
QY 261 Asp 261
Db 803 GAC 805

```

## RESULT 2

```

US-09-610-040-3
; Sequence 3, Application US/09610040
; Patent No. 6465217
; GENERAL INFORMATION:
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Goriach, Jörn
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS
; FILE REFERENCE: 9128.14
; CURRENT APPLICATION NUMBER: US/09/610,040
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 3
; LENGTH: 1006
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-610-040-3
Alignment Scores:
Pred. No.: 8 81e-96 Length: 1006
Score: 818.00 Matches: 182
Percent Similarity: 72.83% Conservatives: 31
Best Local Similarity: 61.13% Mismatches: 68
Query Match: 61.09% Indels: 4
DB: 4 Gaps: 2
US-10-624-061-16 (1-261) x US-09-610-040-3 (1-1006)
QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18
Db 64 ATGGCAAGAGTCTTCGAATCGGATTTCGGGTCTGTGGTTGTTCCAATGTACTGAGTCTTGAC 123
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArg 38
Db 124 TTAATCAGAGATCGTTGATTAGGCAAGAACACCAACGCTCTTCAGCTTGATCAGAGA 183
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58
Db 184 GCTAAGTTTCCACTCAATCTCTCGTTTCGAGGAATCTCGTTGTCTAGATTCTGGAAGT 243
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGluAlaLysAlaGly 78
Db 244 TTCTCTCTCTCAGTGGTTCCTGTCAGAGAGACAGAAATCATCCAAGCTAAGGTAGGA 303
QY 79 ArgTyrLysAsnProGluGlnAlaPhePheProGluAsnLeuProProSerIleVal 98
Db 304 AGATATGAATACCGGAGAGAGAACTCTTCTTCCTTGAGAACATCTCTCACCTCGGTTTT 363
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118
Db 364 CTTACCAATAATCCATCCGCTTTGCAACCCCTAAGGCTCTATCTGTAAACATTAACAAA 423
QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138
Db 424 CAAATCTGGATATTACTTTAAGAAATTCCTCTCTTGTGTCAACCTGGCGATGAT 483
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158
Db 484 GGCAACTATCCATCAACTGCTCTAGTATCTCGCTCTGTTTACAAAGCTCTTTCGAGAAAG 543
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178
Db 544 ATTCACCTCGGTAATTTGTAGCTGAGGTCAAATTCAGAGATGCTCCACAGATTACGAG 603
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198
Db 604 CTTGGCAATTCGGCTCAGGATAGAGAGGCTTTGATGAAGCTGTTCACGCTTTCGAGAAAGTA 663
QY 199 GluGluThrValArgLysArgValGluLysLysAlaValPheGlyGluGluValAsn 218
Db 664 GAAGAAATCGTTTAAAGAGAGAGTGCAGAGAAAGAGAAACGTTTGGACAGAGATGAAA 723
QY 219 LeuAsnSerAspAsnAsnAsnGlnAsnArgLysPhe-----AspProSerValAla 236
Db 724 TTCAACTCTGGCTATGGCGATGAGAGTAAAGAAAGATATAAAGTGAATCCATTCGTTGCC 783
QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValIleTyrLeu 256
Db 784 TCTGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTT 843
QY 257 LeuArgArgLeuAsp 261
Db 844 CTACGCTGCTCGAT 858

```

RESULT 3

US-09-610-040-9/c  
; Sequence 9, Application US/09610040  
; Patent No. 6465217  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Goriach, Jörn  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffmann, Neil  
; APPLICANT: Kloti, Andreas  
; APPLICANT: Zayed, Adel  
; APPLICANT: Ascenzi, Robert  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS  
; FILE REFERENCE: 9128.14  
; CURRENT APPLICATION NUMBER: US/09/610,040  
; CURRENT FILING DATE: 2000-07-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 1006  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-610-040-9  
  
Alignment Scores:  
Pred. No.: 8 81e-96 Length: 1006  
Score: 818.00 Matches: 162  
Percent Similarity: 72.83% Conservative: 31  
Best Local Similarity: 61.13% Mismatches: 68  
Query Match: 61.09% Indels: 4  
DB: 2 Gaps: 2  
  
US-10-624-061-16 (1-261) x US-09-610-040-9 (1-1006)  
QY 1 MetAlaLysAlaAlaGluInSerProAspSerGly-----AsnValTyrThrLeuAla 18  
Db 943 ATGGCAAGACTCTTCGAATCGGATTCGGTTCCTGTTGTTCCTCAATGATCTAGCTCTTGAC 884  
  
QY 19 SerValArgGluAspLeuValArgGluAspThrIleIleTyrGlyLeuIleGluArg 38  
Db 883 TTAATCAGAAATCGTTGATTAGCAGAGACACCATCGTCTTCAGCTTGATCGAGAGA 824  
  
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
Db 823 GCTAAGTTTCCACTCAATCTCCTGCTTCGAGGAATCTCGTGTCTAGATTCTGGAAGT 764  
  
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
Db 763 TTCTCTCTCTCACTAGTTTTTCGTCAGAGACAGAAATCATCAAGCTAGGTAGGA 704  
  
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleVal 98  
Db 703 AGATATGAATACCGGAGAGAAATCCTTTCTCTTGAGAAATCTCTCACTCGGTTTTT 644  
  
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
Db 643 CCTACGCACAAATATCCATCGGCTTTTCACCCCTAAGCTCTATCTGTTAACATTAACAA 584  
  
QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138  
Db 583 CAATCTGGGATATTTACTTTAAGAAATTCCTTCCTTTGTTGTCAACCTCGCGATGAT 524  
  
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158  
Db 523 GGCAACTATCCATCAACTGCTGTAGTATCGCTGTTTACAAAGCTCTTTTCGAGAAGG 464  
  
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
Db 463 ATTCACACGGAATTTAGCTAGGTCAATTCAGAGATGCTCCACAAAGATTACGAG 404  
  
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198

Db 403 CCTCGATTCCGCTCAGGATAGAGAGCCTTTGATGAAGCTGTTGAGCTTTGAGAAAGTA 344  
  
QY 199 GluGluThrValArgLysArgValGluLysAlaValValPheGlyGlnGluValAsn 218  
Db 343 GAAGAAATGTTTAAAGAGAGAGTSCAGAGAAAGCAGAAACGTTTGACAGAGATGATAA 284  
  
QY 219 LeuAsnSerAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236  
Db 283 TTCACTCTGGCTATGCGGATGAGAGTAAGAGAGATATAAGTGGATCCATTCCTTGCC 224  
  
QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
Db 223 TCTCGCATCTACGGGGAATGCTTATCCCTCTCACTAAGCTCGTGGTGGTGGTATCTTT 164  
  
QY 257 LeuArgArgLeuAsp 261  
Db 163 CTACGTCGCTCTCGAT 149  
  
RESULT 4  
US-09-454-279-5  
; Sequence 5, Application US/09454279  
; Patent No. 6627798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: B01299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 5  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (432)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (514)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (516)  
US-09-454-279-5  
  
Alignment Scores:  
Pred. No.: 1 04e-89 Length: 525  
Score: 767.50 Matches: 162  
Percent Similarity: 95.29% Conservative: 0  
Best Local Similarity: 95.29% Mismatches: 7  
Query Match: 57.32% Indels: 4  
DB: 4 Gaps: 0  
  
US-10-624-061-16 (1-261) x US-09-454-279-5 (1-525)  
QY 1 MetAlaLysAlaAlaGluInSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
Db 16 ATGGCCAAAGCAGCAGAAACAAAGTCTGATTCCTGGAATGTGTACACGCTAGCTTCTGTG 75  
  
QY 21 ArgGluAspLeuValArgGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40  
Db 76 AGAGAGGATTTGGTTAGGCAAGAGATACCATCTTTATGGTCTCATTTAGAGAGACCCAA 135  
  
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
Db 136 TTCCTTAGCAATCTTCACACCTATGATGAAAAGTATGCTCAATCCAGGTTTTTTGTGCG 195  
  
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80

Db 196 TCATTGGTGAATTTGTTTAAAGTATACAGAGGCCATTCAGCTAAGCTGGAAGATAC 255  
 QY 81 LysAsnProGluGlnAlaPhePheProGluAsnLeuProSerIleValProSer 100  
 Db 256 AAAAAACCCCTGAAGAAACCCCTTCTCCAGAAAAATTTACCACCATCAATTTGGCCATCT 315  
 QY 101 TyrSerPheGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
 Db 316 TACTCTCTCAACACAGTTTTTGCATCTCTGCTGCTGCTTCAATTAACATTAACAGTCACT 375  
 QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140  
 Db 376 GGGAAATGTAATTTCAAGAGTTACTTCCATTTGCTTACTTCGGGTGATGAGTGAAC 435  
 QY 141 TyrAlaGln-ThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHi 160  
 Db 436 TAATGCGCAAACTGAGCTAATGACCT-TCATTATTGAGTGC-ATCTCTAGAGAT-CA 492  
 QY 160 sTyrGlyLysPheValAlaGluVallys 169  
 Db 493 CTATGCAAGTTTGTAGCTGANGGAA 520

## RESULT 5

US-09-610-040-2  
 ; Sequence 2, Application US/09610040  
 ; Patent No. 6465217  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Boyes, Douglas  
 ; APPLICANT: Davis, Keith  
 ; APPLICANT: Woessner, Jeffrey  
 ; APPLICANT: Gorlach, Jorn  
 ; APPLICANT: Hamilton, Carol  
 ; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Kloti, Andreas  
 ; APPLICANT: Zayed, Adel  
 ; APPLICANT: Ascenzi, Robert  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS  
 ; FILE REFERENCE: 9128.14  
 ; CURRENT FILING DATE: 2000-07-05  
 ; CURRENT APPLICATION NUMBER: US/09/610,040  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1207  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-610-040-2

Alignment Scores:  
 Pred. No.: Length: 6,378-80 1207  
 Score: 696.00 Matches: 131  
 Percent Similarity: 68.08% Conservative: 46  
 Best Local Similarity: 50.38% Mismatches: 75  
 Query Match: 51.98% Indels: 8  
 DB: 4 Gaps: 1  
 US-10-624-061-16 (1-261) x US-09-610-040-2 (1-1207)

QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29  
 Db 235 GATGAGTGTGAGGTTTGTACTCTTGAAGGTATTAGAACTCTTTGATCCGTCAAGAGAC 294  
 QY 30 ThrIleIleTyrGlyLeuLeuGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49  
 Db 295 AGCATATATTGGCTATTGGAGAGCCAGGACTGTTCATAGTCTGATCTATTGAT 354  
 QY 50 GlulysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69  
 Db 355 CCTACTGCTTTTGACATGATGGTTTCAATGCTTTTGGTTGGTGTAGTACATGTTAAAGGC 414  
 QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89

Db 415 ACTGAGAAGCTTCACCTAAGCTTGGTAGTTTAAAGTCTCTGATGACATCTTCTTCTC 474  
 QY 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
 Db 475 CCTGATGATCTACAGAGCTATGTTGCTCTCTCTCAGTACCCAAAGGTGTGCAATTT 534  
 QY 110 GlyAlaAlaSerIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129  
 Db 535 GCTGCTGATTTCGATAAACAATAACAAGAATATGAACATGATCTTACAGACCTTGT 594  
 QY 130 ProLeuLeuAlaThrSerGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
 Db 595 CCAAGACTTGTGAAGAAAGGCGATGATGTAACCTACCGCTCAACAGCTGTCTGTGAGCT 654  
 QY 150 SerLeuLeuGlnSerIleSerArgArgIleHiSerTyrGlyLysPheValAlaGluVallys 169  
 Db 655 ATCTGCTTTCAGTGTCTCTCAAGAGAATCCATTACGGTAAATTTTGTCCAGAAGCTAAA 714  
 QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189  
 Db 715 TTTCAAGCTCACCAGCAGCATACGAGTCCGCCATCAAGACACAGAATAAGGATCGACTG 774  
 QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLys 209  
 Db 775 ATGATATCTGACATCTCCGACTCTGGAAGATGCGATAAAGAGAGAGTTCAGATGAA 834  
 QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAspAsnAsnGluAsnArg 229  
 Db 835 ACCCGAACATACGGCAAGAAGTGAAGTTGGGATGGAGGAGAAAGAGAAGAAGAA 894  
 QY 230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241  
 Db 895 GAAGGGAATGAATCTCATGTTTACAAAATCAGTCCGATCTTAGTTGGTGAATATATGGA 954  
 QY 242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgGluLeuAsp 261  
 Db 955 GATTGATCTGCTCTTACAAAAGAGTTCAAGTGGAGTACTTCTCTCAGAAGACTGGAC 1014

## RESULT 6

US-09-610-040-8/c  
 ; Sequence 8, Application US/09610040  
 ; Patent No. 6465217  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Boyes, Douglas  
 ; APPLICANT: Davis, Keith  
 ; APPLICANT: Woessner, Jeffrey  
 ; APPLICANT: Gorlach, Jorn  
 ; APPLICANT: Hamilton, Carol  
 ; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Kloti, Andreas  
 ; APPLICANT: Zayed, Adel  
 ; APPLICANT: Ascenzi, Robert  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE  
 ; FILE REFERENCE: 9128.14  
 ; CURRENT FILING DATE: 2000-07-05  
 ; CURRENT APPLICATION NUMBER: US/09/610,040  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 8  
 ; LENGTH: 1207  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-610-040-8

Alignment Scores:  
 Pred. No.: Length: 6,378-80 1207  
 Score: 696.00 Matches: 131  
 Percent Similarity: 68.08% Conservative: 46  
 Best Local Similarity: 50.38% Mismatches: 75  
 Query Match: 51.98% Indels: 8  
 DB: 4 Gaps: 1

US-10-624-061-16 (1-261) x US-09-610-040-8 (1-1207)

```
Qy 10 AspSerGlyAsnValThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
Db 973 GATGAGTGTGAGTTGACTCTTGAAGGTATTAGAACTCTTTGATCCGTCAGAGGAC 914
Qy 30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49
Db 913 AGCATTTATTTGGCTATTGGAGAGACCAAGTACTGTTACAAATGCTGTGACTTATGAT 854
Qy 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69
Db 853 CCTACTGCTTTTGACATGATGTTTCAATGCTCTTTGGTTGAGTACATGTTTAAGGC 794
Qy 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89
Db 793 ACTGAGAAGCTTCACGCTAAGTGTGTTAGGTTTAAGAGTCTGATGAACATCCCTTCTTC 734
Qy 90 ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
Db 733 CCTGATGATCTACACAGAGCTATGTTGCTCTCTTCTCAGTACCCCAAGGTGTTGATTTT 674
Qy 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrPheLysMetTyrPheLysGluLeuLeu 129
Db 673 GCTGCTGATTCGTTAAACATAAACAGAGATATCGAACATGTACTTCAGAGACCTTGT 614
Qy 130 ProLeuLeuAlaThrSerGlyAspGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149
Db 613 CCAAGACTGTGAAGAAAGCGCATGATGTAATACTACCGCTCAACAGCTGTCTGTGACGCT 554
Qy 150 SerLeuLeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLys 169
Db 553 ATCTGCTTCAGTGTCTCTCAAGAGATCCATTAAGTAAATTTGTTGCAGAGCTAAA 494
Qy 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGlyLeu 189
Db 493 TTTCAAGCTCACCGAAGCATACGAGTCGCGCATCAAAAGCAAGAAAGATCGACTG 434
Qy 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
Db 433 ATGGATATCTGACATCTCCGACTGGAAGATCGGATAAAGAGAGATTTGAGATGAAA 374
Qy 210 AlaValValPheGlnGluValAsnLeuAsnSerAspAsnAspAsnGluAsnArg 229
Db 373 ACCCGACATACGGCAAGAGTGAAGTTGGATGGAGAGAGAAAGAAAGAAAGAA 314
Qy 230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
Db 313 GAAGGAATGAATCTCATGTTTACAAATCACTCCGATCTTATGTTGGTACTATATGGA 254
Qy 242 AsnTyrValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261
Db 253 GATTGGATATGCTTTAAACAAAGAGGTTCAAGTGAGTACTTGTCTCAGAGACTGGAC 194
```

RESULT 7

```
US-09-454-279-13
; Sequence 13, Application US/09454279
; Patent No. 662798
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/09/454,279
; CURRENT FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1231
```

```
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-454-279-13
```

```
Alignment Scores:
Pred. No.: 1,49e-78 Length: 1231
Score: 685.50 Matches: 135
Percent Similarity: 68.70% Conservatives: 45
Best Local Similarity: 51.53% Mismatches: 81
Query Match: 51.19% Indels: 1
DB: 4 Gaps: 1
```

US-10-624-061-16 (1-261) x US-09-454-279-13 (1-1231)

```
Qy 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20
Db 238 GTACCAAGAGAGAGAGAAACACAGAGATAGATCAAGTGAATACTGACCTTTGGCAACATT 297
Qy 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLys 40
Db 298 AGAACCTCTTGTAGGCAAGAGACAGCATCATATTTCAGCCTCTTAGAGAGACACAG 357
Qy 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60
Db 358 TTTTGTACATCTGATATATATGATAAAATGCTTTCATGTGATGATTTGATGGC 417
Qy 61 SerLeuValGluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80
Db 418 TCTTTGTTGATTCATGTTAGAGAACCGAAACTACATCAACAGGTGGAGATAC 477
Qy 81 LysAsnProGluAsnAlaPheProGluAsnLeuProProSerIleValProSer 100
Db 478 AAGAGCCCTGATGAGCACCCATCTTCTCCGAGAGATCTGCTGAACCACTGTTGCCACCT 537
Qy 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120
Db 538 CTCAGTATCCAAAGGTTTTCATCTCTTTCGATCTTATTAATATCAACAGGAGATT 597
Qy 121 TrpLysMetTyrPheLysGluLeuProLeuLeuAlaThrSerGlyAspAspGlyAsn 140
Db 598 TGGAAATGATTTTGTGATGAGCTTCTTCCAAAGATTAGTGAAGAGAGAGTGTGTAAT 657
Qy 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHis 160
Db 658 TATGATCCAGTGTCTTGTGACAGATCTGCTTGACGGCGCTCTCCAAAGAAATTCAC 717
Qy 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180
Db 718 TATGTTAAGTTTGTGGCAGAGGCTTAAGTTTCAAGAGTCTCTTGAAGCTTACATGCTGCG 777
Qy 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200
Db 778 ATAATAGCAGAGGACTCGATCAACTAATGACCTCTCCTCATGAAACGGTGGAGCGT 837
Qy 201 ThrValArgLysArgValGluLysLysAlaValPheGlyGlnGluValAsnLeuAsn 220
Db 838 GCTATTGAACATAGGTTGGAAGCTTAAGTCTTTGGACAGAGAGGTTGATTTAGGC 897
Qy 221 SerAspAspAsnAspAsnGluAsnArg---LysPheAspProSerValAlaSerSerLeu 239
Db 898 GCTGAAGACACACGCGCTCCACCAATGTACAAAGATAAGGCGCCAGTTTGGTGGCTGAAC 957
Qy 240 TyrLysAsnTyrValIleProLeuThrLysGlnValGlnValGluTyrLeuLeuArgArg 259
Db 958 TACAGCTACAGGATCATCGCGCTTAACCAAGAGGTTCAAGTAGCCTACTTCTGCTGAGGAGA 1017
Qy 260 LeuAsp 261
Db 1018 TTGGAT 1023
```

RESULT 8

```
US-09-454-279-11
; Sequence 11, Application US/09454279
```





150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys 169  
752 ATGTGTTTGCAGATACCTTCAAGAGAAATTCACCTTGGCTAAATTTGTTGCTGAGCCCAAG 811  
170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189  
812 TTTCTGAAATCTCTGCTGCTATGAAACAGCTATCAAGAACCAAGACCGACACAGCTG 871  
190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209  
872 ATGCAACTTCAACCTGCAAGACGGTTGCAAGAGTAGTCAAGAGAGAGAGTTGAGATCAA 931  
210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAspAsnAsp---AsnGluAsn 228  
932 GCCAGAAATTTGGTCAAGACATACAGTAAACGCCACCAAGTGAAGCTGATCTTCC 991  
229 ArgLysPheAspProSerValAlaSerSerLeuTyrLysAsnTrpValIleProLeuThr 248  
992 TACAAATACAACTAGCTTAGTTGCAAACTCTATGAGAAAGGATCATGCCCTCACA 1051  
249 LysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261  
1052 AAGGAGTCCAAATGAGTACTTGGTTAGAGACTGGAT 1090  
RESULT 10  
US-09-610-040-10/c  
; Sequence 10, Application US/09610040  
; Patent No. 6465217  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Goriach, Jörn  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Klotz, Andreas  
; APPLICANT: Zayed, Adel  
; APPLICANT: Ascenzi, Robert  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS  
; FILE REFERENCE: 9128.14  
; CURRENT APPLICATION NUMBER: US/09/610,040  
; CURRENT FILING DATE: 2000-07-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 1217  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-610-040-10  
Alignment Scores:  
Pred. No.: 1,45e-74 Length: 1217  
Score: 654.50 Matches: 128  
Percent Similarity: 67.98% Conservative: 44  
Best Local Similarity: 50.59% Mismatches: 80  
Query Match: 48.88% Indels: 1  
DB: 4 Gaps: 1  
US-10-624-061-16 (1-261) x US-09-610-040-10 (1-1217)  
QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29  
DB 886 GATGAGTGAAGTATTTGAACCTTGAAGCATTTAGACACTCTTTGATGCTCAAGAGGAC 827  
QY 30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49  
DB 826 AGTATTATCTTAATCTTCTGACGAGCTCAGTATCGCTACACGCTGATCTATGAC 767  
QY 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69  
DB 766 GAGGATCCCTTTACTATGGAAGGGTTTCAAGGATCTTTAGTTAGTTTATGGTCAGAGAA 707

QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89  
DB 706 ACTGAAAGCTTCACCGAAGGTGGACAGTACACAGAGCTCTGATGAGCATCCCTTTTC 647  
QY 90 ProGluAsnLeuProSerSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
DB 646 CCACATGCTTGCCTGAACCTATCTCTTCTCCATTCAATACCCACAGGTTTGGCATCGT 597  
QY 110 GlyAlaAlaSerIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129  
DB 586 TGGCCGCAATCGATAAATCAACATCAACAGAGGTGTGGAATATGTATTTCAACACCTTCTC 527  
QY 130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
DB 526 CCCAGACTGCTCAAGCCAGGGGATGACGGTAATTTGGTTTCAGCTCTCTCTGTGACACA 467  
QY 150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys 169  
DB 466 ATGTGTTTGCAGATACCTTCAAGAGAAATTCACCTTGGTAAATTTGTTGCTGAGCCCAAG 407  
QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189  
DB 406 TTTCTGAAATCTCTGCTGCTATGAAACAGCTATCAAGAACCAAGACCGACACAGCTG 347  
QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209  
DB 346 ATGCAACTTCTTAACCTGACGAAACGGTTGAAGAAAGTGTCAAGAGAGAGAGTGTGATCAAA 287  
QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAspAsnAsp---AsnGluAsn 228  
DB 286 GCCAGAAATTTGGTCAAGACATACAGTAAACGCCACCAAGTGAAGCTGATCTTCC 227  
QY 229 ArgLysPheAspProSerValAlaSerSerLeuTyrLysAsnTrpValIleProLeuThr 248  
DB 226 TACAAATACAACTAGCTTAGTTGCAAACTCTATGAGAAAGGATCATGCCCTCACA 167  
QY 249 LysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261  
DB 166 AAGGAGTCCAAATGAGTACTTGGTTAGAGACTGGAT 128  
RESULT 11  
US-09-454-279-17  
; Sequence 17, Application US/09454279  
; Patent No. 6627798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 17  
; LENGTH: 780  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
US-09-454-279-17  
Alignment Scores:  
Pred. No.: 4.3e-68 Length: 780  
Score: 602.00 Matches: 118  
Percent Similarity: 71.08% Conservative: 27  
Best Local Similarity: 57.84% Mismatches: 53  
Query Match: 44.96% Indels: 6  
DB: 4 Gaps: 2  
US-10-624-061-16 (1-261) x US-09-454-279-17 (1-780)  
QY 64 GluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnPro 83

Db 10 GAGTTCTTCGTCGGAAGCCGAGGTCCTGCGACGCAAGGCTGGACACTATCAAAAGCCA 69  
 Qy 84 GluGluAsnAlaPhePheProGluAsnLeuProSerIleValProSerTyrSerPhe 103  
 Db 70 GAAGATTGTCCTCTCTCAAGATCTCCACACCTGCTCTTCTCCACAAAGGTCGC 129  
 Qy 104 LysGlnPheLeuHisProGluAlaAlaSerIleAsnLeuSerIleTyrLysMet 123  
 Db 130 CCAAGGTTTGCACCCCTTTCCTTCATGCTCTGTCGTGATGATGCAATGAAGATG 189  
 Qy 124 TyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsnTyrAlaGln 143  
 Db 190 TATTTCAATGAATCTCTACCAATATTCTCTCGGATGCGATGATGCAACTATGCAGAA 249  
 Qy 144 ThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLys 163  
 Db 250 ACAGTTGCAATAGATTTTGCATGCTGCGAGGCTCTCTCAAGAAATTCATGTGGCAA 309  
 Qy 164 PheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeuIleArgAla 183  
 Db 310 TATGTGTGCTGAGGTGAATTCAGAGACGCGCTCAAGATTATAGCCACCAATACGTGCT 369  
 Qy 184 LysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGluThrValArg 203  
 Db 370 AAGGACACTAATGCTCTCTGATGAATCTTAACTGCTGCGGCTGTGAGAAAGGTCAAG 429  
 Qy 204 LysArgValGluLysLysAlaValPheGlyGlnGluValAsnLeu----- 219  
 Db 430 AAGAGAGTAGAAGAGGCAAGGATATTTGACAGAGATGTCCTCTGGAGGACAGGTGA 489  
 Qy 220 AsnSerAspAspAsnAsp-----AsnGluAsnArgLysPheAspProSerValAlaSer 237  
 Db 490 GCGAAGCAAGATGGTGATGCTGTCGACAGTCACTGTAAGATGATCCAAAGTCTTCT 549  
 Qy 238 SerLeuTyrLysAsnTyrPheProLeuThrLysGluValGlnValGluTyrLeuLeu 257  
 Db 550 AAGCTATATGATATGCTGGTAATGCCCTTAAAGCAAGGATGTTGAAGTCGAATATCTTC 609  
 Qy 258 ArgArgLeuAsp 261  
 Db 610 CGCGCTCTTGAC 621

## RESULT 12

US-09-610-040-6  
 ; Sequence 6, Application US/09610040  
 ; Patent No. 6465217

## GENERAL INFORMATION:

; APPLICANT: Boyes, Douglas  
 ; APPLICANT: Davis, Keith  
 ; APPLICANT: Woessner, Jeffrey  
 ; APPLICANT: Gorchach, John  
 ; APPLICANT: Hamilton, Carol  
 ; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Klotz, Andreas  
 ; APPLICANT: Zayed, Adel  
 ; APPLICANT: Ascenzi, Robert  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS  
 ; FILE REFERENCE: 9128.14  
 ; CURRENT APPLICATION NUMBER: US/09/610,040  
 ; CURRENT FILING DATE: 2000-07-05  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 6  
 ; LENGTH: 5176  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana

US-09-610-040-6

Alignment Scores:  
 Pred. No.: 7,07e-65 Length: 5176  
 Score: 587.00 Matches: 162

Percent Similarity: 36.50% Conservative: 30  
 Best Local Similarity: 30.80% Mismatches: 69  
 Query Match: 43.84% Indels: 268  
 DB: 4 Gaps: 6  
 US-10-624-061-16 (1-261) x US-09-610-040-6 (1-5176)  
 Qy 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 Db 2542 ATGGCAAGAGTCTTCAATCGGATTCGGGTTCGTGTGTCCATGTACTGTAGTCTTGAC 2601  
 Qy 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArg 38  
 Db 2602 TTAATCAGAGAATCGTTGATTAGGCAAGAAGACACCATCGTCTTCAGCTTGATCGAGAGA 2661  
 Qy 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 Db 2662 GCTAAGTTTCCACTCAATTCCTCTGCTTTCGAGGAATCTCGTTGTCTAGATTCTCGAAGT 2721  
 Qy 59 CysGlySerLeuValGluPheValLysAsnThrGluAlaIleGlnAlaLys----- 76  
 Db 2722 TTCTCTCTCTCACTGAGTTTTCGTCAGAGAGACAGAAATCATCCAGCTAAGTTTGC 2781  
 Qy 76 ----- 76  
 Db 2782 TTCCCAATTTAAAACTGATCCTTTTGTCTAAATPAGATACAGAGATATCAATGCTCGT 2841  
 Qy 76 ----- 76  
 Db 2842 TTGATTCGGTTTTGGTATAGCATTTGTTTAGATTGTTCCATGAATTAGCAGAAAGTAAG 2901  
 Qy 76 ----- 76  
 Db 2902 CTACAAGTCAACTTCATTCAGGTTTTTAATAAGCCTGGATTCTTGAAATTAGCATGCTTTT 2961  
 Qy 76 ----- 76  
 Db 2962 GTTTCGTATGTCTCTCCATTCGAAAGATGATACTTCGGCTTTCGCTGTATAATCTC 3021  
 Qy 76 ----- 76  
 Db 3022 ATTGTGTGATAAATCTCTGTTTGTGATTGAGTGGCAATCTCCCAATAAAGGCTCCGACT 3081  
 Qy 76 ----- 76  
 Db 3082 TTATCATATGATACAGAGATTTCCTTATGAAAACCTCATTAATATGTGGAGATTGGAATG 3141  
 Qy 76 ----- 76  
 Db 3142 GAGGACTATTGTTTCTATTATTTTATAATGTCTGAAAGTCTTATTTCATTAATATTCAT 3201  
 Qy 76 ----- 76  
 Db 3202 CTCATGTTTATATCTTAAGTTTCTGGATATTGAGCCTATATGTTTGTTCATTTGTTT 3261  
 Qy 77 -----AlaGlyArgTyrLysAsnProGlu 84  
 Db 3262 ACTTGAAACCTTATGTGTATGTATATATATATAGTAGGAGAGATATGATACCCGGA 3321  
 Qy 85 GluAsnAlaPhePheProGluAsnLeuProSerIleValProSerTyrSerPheLys 104  
 Db 3322 GAGAATCCTTTCTTCTCTTGAGAACATTCCTCACTCGGTTTTTCTACGCAAAATATCCA 3381  
 Qy 104 ----- 104  
 Db 3382 TCGGTATGATGATAGTAAGTCTTGACATTTTCTTAGACTCTGATGCTTAGTCTA 3441  
 Qy 105 -----GlnPheLeuHisProGlyAlaAlaSerIleAsnIle 116  
 Db 3442 ACAGTACTTTTCTTCTCTTGTGATTAGGCT-TTGACCCCTAAGGCTCTATCTGTTAAT 3500  
 Qy 117 AsnLysSerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGly 136

3501	Db	AA	CAAA	CAAA	CTGGGA	ATATTACTTTAAAGAA	ATGCTCTCTTTGTTGTTC	AAACCTGGC	3560
137	QY	ASP	AS	PG	LY	ASN	TYR	VAL	154
3561	Db	GAT	GAT	GCGCA	ACTATCCAT	CAACTGCTGCTAGT	GATCTCGCCTGTTTACA	AGTAAAGGAG	3619
154	QY	---	---	---	---	---	---	---	154
3620	Db	AT	GAT	TGAGT	ATACAT	CAAAATCAGCTCTACTTTTGGCTAA	TGATGTCGTATCTGAT	CTGATA	3679
155	QY	---	---	---	---	---	---	---	167
3680	Db	TG	TTG	TGAT	CTTGCT	TAAGGCTCTTTTCGAGAAGG	ATTCACCTACGTA	AAATTTG	3739
168	QY	VAL	LYS	PHE	ARG	ASP	ALA	PRO	184
3740	Db	GT	CAA	ATTCAGAG	ATGCTCCACA	AGATTACGAGCCTCGGATTC	CGGCTCA	GGTAA	3798
184	QY	---	---	---	---	---	---	---	184
3799	Db	AGT	GT	CACA	ATTGTG	GATTCTGTTCTTTCACTG	TGGTTTTAAATGAT	TATGATTCACACCA	3858
185	QY	---	---	---	---	---	---	---	197
3859	Db	TCG	TTG	ATTTTCGAC	TGTAGGATAGAGAG	CGCTTTGATGAAGCTGTTGAC	GGTTTGAGAAA	3918	
198	QY	VAL	GLU	GLU	THR	VAL	ARG	LYS	217
3919	Db	GT	AGA	GAATGGTTA	AGAAGAGAGAGTCCAGA	AGAAAGCAGTAAAGCTTTG	GCACAAAGTA	3978	
218	QY	ASN	LEU	ASN	SER	ASP	ASP	ASN	235
3979	Db	AAAT	TCA	CTCTG	CTCTG	CGATGCGGATGAGAGT	TAAGAAGAGATATAAAGT	GTGGATTC	4038
236	QY	ALA	SER	SER	LEU	TYR	LYS	ASN	255
4039	Db	GCCT	CTCGCAT	CTACGGG	GAATGCGTTATCCCTCTCA	CTAAAGCTCGTTGAGGTT	GAGTAT	4098	
256	QY	LEU	LEU	ARG	ARG	LEU	ASP	261	
4099	Db	CTT	CTACG	TCGCT	CTCGAT	4116			

RESULT 13  
 US-09-610-040-5  
 ; Sequence 5, Application US/09610040  
 ; Patent No. 6465217  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Boyes, Douglas  
 ; APPLICANT: Davis, Keith  
 ; APPLICANT: Woessner, Jeffrey  
 ; APPLICANT: Goxlach, Jörn  
 ; APPLICANT: Hamilton, Carol  
 ; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Kloti, Andreas  
 ; APPLICANT: Zayed, Adel  
 ; APPLICANT: Ascenzi, Robert  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATASE  
 ; TITLE OF INVENTION: CHORISMATASE MUTASE EXPRESSION OR ACTIVITY IN PLANTS  
 ; FILE REFERENCE: 9128.14  
 ; CURRENT APPLICATION NUMBER: US/09/610,040  
 ; CURRENT FILING DATE: 2000-07-05  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 5  
 ; LENGTH: 5099  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (4049)..(4049)  
 ; OTHER INFORMATION: "n" indicates any nucleotide (A, C, T, or G).  
 US-09-610-040-5

```

QY 185 -----AspLysGluGlyLeuMetLysLeuThrPheThr 197
Db 4369 ACATGGCAACATACGATTGCAGGATAAGGATGCTGATGATGCTGACATTCGGA 4428
QY 197 ex-ValGluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGluGlu 217
Db 4429 CTTGTGAAGATCCGATAAAGAGAGAGTTGATGATGAAACCCGACATACCGGCAAGAAG 4488
QY 217 alAsnLeuAsnSerAspAspAsnAspAsnGluAsnArg----- 229
Db 4489 TCAAAAGTTGGGATGGAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4548
QY 230 --LysPheAspProSerValAlaSerSerLeuTyrLysAsnTyrValIleProLeuThrL 249
Db 4549 ACAAAATCAGTCCGATCTAGTTGTGATCTATATGGAGATTGGATCAGCTTTAAACA 4608
QY 249 ysGluValGlnValGluTyrLysLeuArgArgLeuAsp 261
Db 4609 AAGAGGTTCAAGTGGAGTACTTCTCAGAGAGACTGGAC 4646

RESULT 14
US-09-454-279-7
; Sequence 7, Application US/09454279
; Patent No. 6627798
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/09/454,279
; CURRENT FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (417)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (484)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (536)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (551)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (565)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (577)
US-09-454-279-7
Alignment Scores:
Pred. No.: 5,62e-58 Length: 579
Score: 522.00 Matches: 103
Percent Similarity: 67.19% Conservative: 26
Best Local Similarity: 53.65% Mismatches: 57
Query Match: 38.98% Indels: 6
DB: 4 Gaps: 2

US-10-624-061-16 (1-261) x US-09-454-279-7 (1-579)
QY 64 GluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnPro 83

```

```

Db 3 GAGTTCTTCTGTCGGGAAGCCGAGGTCCTGCACGCCAGGCTGGCACTATCAAAAGCCA 62
QY 84 GluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSerTyrSerPhe 103
Db 63 GAAGATGTTCCATCTCTCTCAGATCTTCCACACCTGCTCTTCTTACCAAGGTCGC 122
QY 104 LysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIleTrpLysMet 123
Db 123 CCAAAGGTTTTCACCCCTTTTCTTCATGCTGCTGATGATGATGATGATGATGATGATG 182
QY 124 TyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGln 143
Db 183 TATTTCAATGAATGCTACCATTTACTGCGGATGGCGATGATGCGCAACTATCAGAA 242
QY 144 ThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArgIleHisTyrGlyLys 163
Db 243 ACAGTTGCATTAGATTTTGCATGCTCTGCAGGCTCTCTCAAGAAGAATTCATTTGTCGCAA 302
QY 164 PheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeuIleArgAla 183
Db 303 TATGTTGCTGAGGTGAATTTCAAGACCGCGCTCAAGATTATAGCCCACTATACGTGCT 362
QY 184 LysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGluThrValArg 203
Db 363 AAGGACACTAATGCTCTGATGAATTTACTAAGCTTCAGGCTGTTGAAGAAAAGTCAAG 422
QY 204 LysArgValGluLysLysAlaValValPheGlyGlnGluValAshLeu----- 219
Db 423 AAGAGAGTAGAGAAGAAGCAAGGATATTTGGACAGAATGTCACTCTGGAGGACAGTGTA 482
QY 220 AsnSerAspAspAsnAsp-----AsnGluAsnArgLysPheAspProSerValAlaSer 237
Db 483 GCGAAGCAAGATGGTGATGCTGTCGACAGTCACTGTAAGTTGATTCCAAAGTGTCTTCT 542
QY 238 SerLeuTyrLysAsnTyrValIleProLeuThrLys 249
Db 543 AAGCTATATGATATGCGGGAATGCCCTTTAAACGAA 578

RESULT 15
US-09-454-279-3
; Sequence 3, Application US/09454279
; Patent No. 6627798
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/09/454,279
; CURRENT FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (421)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (435)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (459)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (490)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (507)

```

```

FEATURE:
; NAME/KEY: unsure
; LOCATION: (526)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (547)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (566)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (579) ..(580)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (582)
US-09-454-279-3

Alignment Scores:
Pred. No.: 2,93e-09 . Length: 601
Score: 144.50 Matches: 41
Percent Similarity: 59.43% Conservative: 22
Best Local Similarity: 38.68% Mismatches: 34
Query Match: 10.79% Indels: 10
DB: 4 Gaps: 2

US-10-624-061-16 (1-261) x US-09-454-279-3 (1-601)
QY 1 MetAlaLysAlaAlaGluClnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20
DB 231 GTGACCAAGGAGGAGAAACAGAGGATAGATCAAAGTGAATACACCTTGGACACATT 290
QY 21 ArgGluAspLeuValArgGlnGlnAspThrIleTyrGlyLeuIleGluArgAlaLys 40
DB 291 AGAACCTCCTTGGTTAGGCAAGACAGACATCATATTGAGCCTCTTAGAGAGACACAG 350
QY 41 -PheProSerAsnSerHisThrTyrAspClnLys-TyrAlaClnIleGlnGlyPheCysG 60
DB 351 TTTTGCTACAAATGCTGATATATATGATAAAATGCTTCCATGTTGGATGGATTGATG 410
QY 60 LysSerLeuValGluPheValVal--LysAsnThrGluAlaIleGlnAlaLysAlaGlyA 79
DB 411 GCCTTGGTTNAAATTCATGGTTANGAGAAACCGAATAACTACATCACAAGTTTGGGA 470
QY 79 IGTyrlLysAsnProGluGluAsnAlaPhePheProGluAsnLeuPro----- 94
DB 471 GATACAAAGAGCCCTGATGANCACCA-TTCTTTCCGGANGATCTGCCTGAACAATGTNGCA 529
QY 95 -----ProSerIle 97
DB 530 ACCTCCCCAGTATC 543

```

Search completed: September 25, 2004, 05:06:03  
Job time : 112 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 25, 2004, 04:17:54 ; Search time 514 Seconds  
(without alignments)

2571.891 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKAAEQPSDGNVTLASV.....NWVPLTRKQVQVLLRLD 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10624061/runat\_24092004\_092409\_17014/app\_query.fasta\_1.455  
-DB=published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS-human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MTN=0 -ALIGN=15 -MODES=LOCAL -OUTFMT=pct -NCRM=ext -HEARSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10624061@cgn 1 1 537 -runat\_24092004\_092409\_17014  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq2:  
14: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:  
17: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:  
18: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:  
19: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

ALIGNMENTS

RESULT 1

US-09-454-279-15  
; Sequence 15, Application US/09454279  
; Publication No. US20020184658A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omelayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: B01299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 15  
; LENGTH: 1020  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-454-279-15

Alignment Scores:

Pred. No.: 3 54e-164 Length: 1020  
 Score: 1339.00 Matches: 261  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-624-061-16 (1-261) x US-09-454-279-15 (1-1020)

QY 1 MetAlaLysAlaAlaGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
 DB 23 ATGGCCAAAGCAGCAGAAAGAGTCTGATTCGGGAATGTGTACACGCTAGCTTCTGTG 82  
 QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGlnGlyAlaLys 40  
 DB 83 AGAGAGGATTTGGTTAGGCAAGAGGATACCATTTATGGTCTCATTTGAGAGGCCAAG 142  
 QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
 DB 143 TTCCCTAGCAATTTCTCACACCTATGATGAAAGTATGCTCAATCCAGGTTTTTGTGGC 202  
 QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
 DB 203 TCATTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAGCTAAGCTGGAAGATAC 262  
 QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100  
 DB 263 AAAAACCCCTGAAGAAAGCCCTTCTCCAGAAATTTTACCACCATCAATTTGCCATCT 322  
 QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
 DB 323 TACTCCTTCAACAGATTTTGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 382  
 QY 121 TrpLysMetTyrPheLysGlnLysLeuProLysLeuAlaThrSerGlyAspGlyAsn 140  
 DB 383 TGGAAATGTATTTCAAGAGATTTGATGAAATTTTCAATTAACATTAACAGGATTCAC 442  
 QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArgIleHis 160  
 DB 443 TATGGCAAACTGCAAGTAAATTCAGGAGTCTCTCAAGACTACAGCCCTTTA 502  
 QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
 DB 503 TATGAAAGTTTGTAGCTGAGTGAAATTCAGGATGCTCTCAAGATTAACAGGCTTTA 562  
 QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220  
 DB 623 ACGGTGAGAAAGAGATTTGAAAGAGGCTGTGGTGTGGGAGGAGGATCTTAAC 682  
 QY 221 SerAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240  
 DB 683 AGTGATGACATGACATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTGTAC 742  
 QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeu 260  
 DB 743 AAAAATGGGTGATACCTCTCACCAGAGGTTTCAAGTTGAGTACCTCTTTCGCGCGTCTA 802  
 QY 261 Asp 261  
 DB 803 GAC 805

## RESULT 2

US-10-624-061-15  
 ; Sequence 15, Application US/10624061  
 ; Publication No. US20040019929A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Falco, S. Carl  
 ; APPLICANT: Famedu, Omolayo O.  
 ; APPLICANT: Lee, Jian-Wing

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 ; FILE REFERENCE: B01299 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/624,061  
 ; CURRENT FILING DATE: 2003-07-21  
 ; PRIOR APPLICATION NUMBER: US/09/454,279

; PRIOR FILING DATE: 1999-12-03  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 15  
 ; LENGTH: 1020  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; US-10-624-061-15

## Alignment Scores:

Pred. No.: 3 54e-164 Length: 1020  
 Score: 1339.00 Matches: 261  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 16 Gaps: 0

US-10-624-061-16 (1-261) x US-10-624-061-15 (1-1020)

QY 1 MetAlaLysAlaAlaGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
 DB 23 ATGGCCAAAGCAGCAGAAAGTCTGATTCGGGAATGTGTACACGCTAGCTTCTGTG 82  
 QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGlnGlyAlaLys 40  
 DB 83 AGAGAGGATTTGGTTAGGCAAGAGGATACCATTTATGGTCTCATTTGAGAGGCCAAG 142  
 QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
 DB 143 TTCCCTAGCAATTTCTCACACCTATGATGAAAGTATGCTCAATCCAGGTTTTTGTGGC 202  
 QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
 DB 203 TCATTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAGCTAAGCTGGAAGATAC 262  
 QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100  
 DB 263 AAAAACCCCTGAAGAAAGCCCTTCTCCAGAAATTTTACCACCATCAATTTGTGCCATCT 322  
 QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
 DB 323 TACTCCTTCAACAGATTTTGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 382  
 QY 121 TrpLysMetTyrPheLysGlnLysLeuProLysLeuAlaThrSerGlyAspGlyAsn 140  
 DB 383 TGGAAATGTATTTCAAGAGATTTGATGAAATTTTCAATTAACATTAACAGGATTCAC 442  
 QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArgIleHis 160  
 DB 443 TATGGCAAACTGCAAGTAAATTCAGGAGTCTCTCAAGACTACAGCCCTTTA 502  
 QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
 DB 503 TATGAAAGTTTGTAGCTGAGTGAAATTCAGGATGCTCTCAAGATTAACAGGCTTTA 562  
 QY 181 IleArgAlaLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 200  
 DB 563 ATTCGAGCTAAGGATTAAGAGAGGATTTGATGAAATTTTGGATTTTACAGGCTTTAAGAG 622  
 QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220  
 DB 623 ACGGTGAGAAAGAGATTTGAAAGAGGCTGTGGTGTGGGAGGAGGATCTTAAC 682  
 QY 221 SerAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240  
 DB 683 AGTGATGACATGACATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTGTAC 742  
 QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeu 260  
 DB 743 AAAAATGGGTGATACCTCTCACCAGAGGTTTCAAGTTGAGTACCTCTTTCGCGCGTCTA 742



QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgA:GLeu 260  
DB 743 AAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGTTGAGTACCTCTTGGCCGCTCTA 802  
QY 261 Asp 261  
DB 803 GAC 805  
RESULT 3  
US-10-425-114-7639  
; Sequence 7639, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 7639  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700677130\_FLI  
US-10-425-114-7639  
Alignment Scores:  
Pred. No.: 3,86e-164 Length: 1079  
Score: 1339.00 Matches: 261  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 13  
US-10-624-061-16 (1-261) x US-10-425-114-7639 (1-1079)  
QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
DB 110 ATGCCCAAGAGCAGACAAAGCTCTGATCTGGGAATGTGTACACGCTAGCTCTGTG 169  
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40  
DB 170 AGAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGCTCTCATTTGAGAGAGCCAA 229  
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
DB 230 TTCCTAGCAATTCACACCTATGATGAAAGATGTCTCAAAATCCAGGGTTTTTGTGCG 289  
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
DB 290 TCATTGTGGAAATTTGTTTGAAGATACAGAGCCATTCAAGCTAAGCTGGAGATAC 349  
QY 81 LysAsnProGluGlnAsnAlaPhePheProGluAsnLeuProPheSerIleValProSer 100  
DB 350 AAAACCCCTGAAGAAAACGGCTCTTCCCAAGAAATTTACCAATCAATTTGGCCATCT 409  
QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
DB 410 TACTCTTCAACAGATTTTTCATCCCTGGTGTCTTCAATTAACATAAACAAGTCCATC 469  
QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsn 140  
DB 470 TGGAAAATGATTTCAAGAGATTTACTTCCATTCCTGCTACTTCCGGGTGATGCGCAAC 529  
QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgIleHis 160

DB 530 TATGCCAAATCGCAGCTAATGACCTTTCATTTATTCAGTCCATCTCTAGAAGGATTCAC 589  
QY 161 TyrGlyLysPheValAlaGluValLysPheAspAlaProGlnAspTyrGluProLeu 180  
DB 590 TATGGAAGTTTGTAGCTGAGGTGAATTCAGGATGCTCTCAAGACTACGAGCCCTTTA 649  
QY 191 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200  
DB 650 ATTCGAGCTAAGGATAAAGAGGATTTGATGAAATTTGTCATTTTACAAGCGTTGAAGAG 709  
QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220  
DB 710 ACGGTGAGGAGAGAGTTGAAAGAGCGCTGTGCTGTTTGGCGGAGGATGAATCTTAAC 769  
QY 221 SerAspAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240  
DB 770 AGTGATGACCAATGACCAATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTGTAC 829  
QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgGlu 260  
DB 830 AAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGTTGAGTACCTCTTGGCCGCTCTA 889  
QY 261 Asp 261  
DB 890 GAC 892  
RESULT 4  
US-10-424-599-93036  
; Sequence 93036, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 93036  
; LENGTH: 1653  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_55025C.1  
US-10-424-599-93036  
Alignment Scores:  
Pred. No.: 1.01e-160 Length: 1653  
Score: 1315.00 Matches: 255  
Percent Similarity: 100.00% Conservative: 6  
Best Local Similarity: 97.70% Mismatches: 0  
Query Match: 98.21% Indels: 0  
Gaps: 13  
US-10-624-061-16 (1-261) x US-10-424-599-93036 (1-1653)  
QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
DB 447 ATGCCCAAGCAGACAAAGCTCTGATTTGGGAATGTGTACACGCTAGCTCTGTG 506  
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40  
DB 507 AGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTATGCTCTCATTCAGAGAGCCAA 566  
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
DB 567 TTCCTAGCAATTCACACCTATGATGAAAGATGTCTCAAAATCCAGGGTTTTTGTGCG 626  
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80

```

Db      627 TCATTCGTAGAATTTCGTGTAAGATAACAGAGGCCATTCAAGCTGAAGATAC 686
QY      81  LysAsnProGluGluAsnAlaPheProGluAsnLeuProSerIleValProSer 100
Db      687 AAAAACCCCTGAAGAAACCCCTCTTCCAGAAAAATTTACCAACCAATCAATTGTGCCATCT 746
QY      101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120
Db      747 TACTCTCTCAACACAGTTTTTGCATCTCGAGCTGCTTCAATTAACATTACACAGTCCATC 806
QY      121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsn 140
Db      807 TGGAAATATATTTCCAGAGTTACTTCCATTGCTTGTCTTCTCGGGTGATGGAAC 866
QY      141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArgIleHis 160
Db      867 TATGCACAACCTGCAGCTAATGATCTTTCAATTTTGCAGGCCATCTCTAGAGGATTCAC 926
QY      161 TyrGlyPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180
Db      927 TATGGAAGTTTGTAGCTGAGTGAAATTCAGGATGCTCTCAAGACTACGAGCCCTTA 986
QY      181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200
Db      987 ATTCGAGCTAAGGATAGAGGATGATGAATTTGACATTTACAGCGTTGAGAG 1046
QY      201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220
Db      1047 ACGGTGAGAGAGAGGTTGAAAGAGGCTGTGGTGTGTTGGCGAGGAAGTCAATCTTGAC 1106
QY      221 SerAspAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerIleTyr 240
Db      1107 AATGAGGAGATGATGAATGAACCGTAATTTGATCCATCAGTGGCTTCTAGCTTGATC 1166
QY      241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuLeuArgLeu 260
Db      1167 AAAAATGGGTGATACCCCTTACCAAGGAGGTTGAGTTGAGTACTTATGGCGGCTA 1226
QY      261 Asp 261
Db      1227 GAC 1229

```

## RESULT 5

US-10-425-114-15059  
 ; Sequence 15059, Application US/10425114  
 ; Publication No. US20040034888A1

## GENERAL INFORMATION:

APPLICANT: Liu, Jingdong  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Screen, Steven E  
 APPLICANT: Tabaska, Jack E  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 15059

LENGTH: 992

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: LIB3050-017-E9\_FLI  
 US-10-425-114-15059

## Alignment Scores:

Pred. No.: 4.51e-116 Length: 992  
 Score: 970.00 Matches: 187  
 Percent Similarity: 97.99% Conservative: 8  
 Best Local Similarity: 93.97% Mismatches: 4

Query Match: 72.44% Indels: 0  
 DB: 13 Gaps: 0

US-10-624-061-16 (1-261) x US-10-425-114-15059 (1-992)

```

QY      63 ValGluPheValLysLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsn 82
Db      2 GTAGATTGTTCTTGAAGATACAGAGGCCATTCAAGCTAAGCTGGAAGATACAAAC 61
QY      83 ProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSerTyrSer 102
Db      62 CCTGAAGAAACCCCTCTTCCAGAAAAATTTACCAACCATCAATTTGTGCGCATCTTACTCC 121
QY      103 PheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIleTrpLys 122
Db      122 TTCAACAGTTTTTGCATCTCTGGAGCTGCTTCAATTAACATTAAACAGTCCATCTGGAAA 181
QY      123 MetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAla 142
Db      182 ATGTATTTCCAGAGTTACTTCCATTGGTTGCTACTTTCGGGGGATGATGGAACATATGCA 241
QY      143 GlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArgIleHisTyrGly 162
Db      242 CAAACTGCAGCTAATGATCTTTCAATTTGCGGGCCATCTCTAGAGGATTCACATATGGA 301
QY      163 LysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeuIleArg 182
Db      302 AAGTTTGTAGCTGAGTGAAATTCAGGATGCTCTCAAGACTACGAGCCTTTAATTGCA 361
QY      183 AlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGluThrVal 202
Db      362 GCTAAGGATTAAGAGGATTGATGAATTTGACATTTTACAAGCGTTTCAAGAGACGGTG 421
QY      203 ArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsnSer 222
Db      422 AGGAAGAGGTTGAAAGAGGCCCGCAGTGTTTGGCGAGGAAGTGAGTCTTGCAACATGAG 481
QY      223 AspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyrLysAsn 242
Db      482 GACGATGATGAGGAAACCATTAAGTTTGTATCCATCAGTGGCTTCTAGCTTGTACAAAAT 541
QY      243 TrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeuAsp 261
Db      542 TGGGTGATACCCCTTACCAAGAGGTTTCAGTTGAGTACTATTGGCGGCTTAGAC 598

```

## RESULT 6

US-09-938-842A-972  
 ; Sequence 972, Application US/09938842A  
 ; Patent No. US20020160378A1

## GENERAL INFORMATION:

APPLICANT: Harper, Jeff  
 APPLICANT: Kreps, Joel  
 APPLICANT: Wang, Xun  
 APPLICANT: Zhu, Tong  
 TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 TITLE OF INVENTION: SAME, AND METHODS OF USE  
 FILE REFERENCE: SCRI3300-3  
 CURRENT APPLICATION NUMBER: US/09/938,842A  
 CURRENT FILING DATE: 2001-08-24  
 PRIOR APPLICATION NUMBER: US 60/227,866  
 PRIOR FILING DATE: 2000-08-24  
 PRIOR APPLICATION NUMBER: US 60/264,647  
 PRIOR FILING DATE: 2001-01-16  
 PRIOR APPLICATION NUMBER: US 60/300,111  
 PRIOR FILING DATE: 2001-06-22  
 NUMBER OF SEQ ID NOS: 5379  
 SEQ ID NO 972

LENGTH: 798

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-938-842A-972

Alignment Scores:

Pred. No.: 2,14e-96 Length: 798  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservative: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 9 Gaps: 2

US-10-624-061-16 (1-261) x US-09-938-842A-972 (1-798)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSergly-----AsnValTyrThrLeuAla 18  
 Db 1 ARGCAAGAGCTTCGAATCGGATTCGGTTCCTGTTTCCATGCTACTGAGCTTGAC 60  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuLeuGluArg 38  
 Db 61 TTAATCAGAGATCGTTGATTAGGCAAGACACCATCGTCTTCACGCTTGCATCGAGAGA 120  
 QY 39 AlalysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 59  
 Db 121 GCTAAGTTCACATCAATTCCTCGTTCGAGGAATCTCGTTGTCFAGATTCTGGAAGT 180  
 QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 Db 181 TTCTCTCTCTCAGTCTGTTTTCGTCAGAGACAGAAATCATCCAGCTTAGGTAGGA 240  
 QY 79 ArgTyrLysAsnProGluGlnAlaPhePheProGluAsnLeuProSerIleVal 98  
 Db 241 AGATATGAATACCCGGAAGAAATCCCTTCTTCCTCGAGAACATCTCTCACTCGGTTT 300  
 QY 99 ProSerTyrSerPheLysGlnPheLysHisProGlyAlaAlaSerIleAsnLys 118  
 Db 301 CCTACGCACAAATATCCATCGCTTTCACCCCTAAGCTCTAICTGTATACATTAACAA 360  
 QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138  
 Db 361 CAATATCGGATATTAATTTAAGAAATTCCTTCCTGTTTGTCAACCTGCGGATGAT 420  
 QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 Db 481 ATTCACATCGGTAATTTAGCTGAGGCTCAATTCAGAGATGCTCCCAAGATTACGAG 540  
 QY 237 SerSerLeuTyrLysAsnTyrValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
 Db 721 TCTCGCATCTACGGGAATGGCTTATCCCTCACTAAGCTCGTTGAGGTGAGTATCTT 780  
 QY 257 LeuArgArgLeuAsp 261  
 Db 781 CTACGTCGCTTCGAT 795

# RESULT 7

US-09-938-842A-972  
 ; Sequence 972, Application US/09938842A  
 ; Publication No. US20040009476A9  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong  
 TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 FILE REFERENCE: SAME, AND METHODS OF USE  
 CURRENT APPLICATION NUMBER: US/09/938,842A  
 PRIOR FILING DATE: 2001-08-24  
 PRIOR APPLICATION NUMBER: US 60/227,866  
 PRIOR FILING DATE: 2000-08-24  
 PRIOR APPLICATION NUMBER: US 60/264,647  
 PRIOR FILING DATE: 2001-01-16  
 PRIOR APPLICATION NUMBER: US 60/300,111  
 PRIOR FILING DATE: 2001-06-22  
 NUMBER OF SEQ ID NOS: 5379  
 SEQ ID NO 972  
 LENGTH: 798  
 TYPE: DNA  
 ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-972

Alignment Scores:  
 Pred. No.: 2,14e-96 Length: 798  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservative: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 9 Gaps: 2

US-10-624-061-16 (1-261) x US-09-938-842A-972 (1-798)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSergly-----AsnValTyrThrLeuAla 18  
 Db 1 ATGCAGAGCTTCGAATCGGATTCGGTTCCTGTTTCCATGCTACTGAGCTTGAC 60  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuLeuGluArg 38  
 Db 61 TTAATCAGAGATCGTTGATTAGGCAAGACACCATCGTCTTCACGCTTGCATCGAGAGA 120  
 QY 39 AlalysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 Db 121 GCTAAGTTCACATCAATTCCTCGTTCGAGGAATCTCGTTGTCTAGATCTGGAAGT 180  
 QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 Db 181 TTCTCTCTCTCAGTCTGTTTTCGTCAGAGACAGAAATCATCCAGCTTAGGTAGGA 240  
 QY 79 ArgTyrLysAsnProGluGlnAlaPhePheProGluAsnLeuProSerIleVal 98  
 Db 241 AGATATGAATACCCGGAAGAAATTCCTTCCTGTTGAGAACATTCCTCACTCGGTTT 300  
 QY 99 ProSerTyrSerPheLysGlnPheLysHisProGlyAlaAlaSerIleAsnLys 118  
 Db 301 CCTACGCACAAATATCCATCGGCTTTCACCCCTAAGCTCTAICTGTATACATTAACAA 360  
 QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138  
 Db 361 CAATATCGGATATTAATTTAAGAAATTCCTTCCTGTTTGTCAACCTGCGGATGAT 420  
 QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArg 158  
 Db 421 GGCAACTATCCATCAATCGCTGCTAGTATCTCGCTGTTTACAGCTCTCTTACCATTAACAA 480  
 QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 Db 481 ATTCACATCGGTAATTTAGCTGAGGCTCAATTCAGAGATGCTCCCAAGATTACGAG 540  
 QY 179 ProLeuIleArgAlaLysAspLysGluLysLeuMetLysLeuLeuThrPheThrSerVal 198  
 Db 541 CCTCGCATCTACGGGAATGGCTTATCCCTCACTAAGCTCGTTGAGGTGAGTATCTT 600  
 QY 199 GluGluThrValArgLysArgValGluLysLysAlaValPheGlyGlnGlnValAsn 218  
 Db 601 GAAGAAATGGTTAAGAAAGAGAGTGCAGAGAAAGACGTTTGGACAAAGTAAATA 660

QY 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236  
 Db 661 TTCAATCTGGCTATGGCGATGAGATAGAGAAGTATAAAGTGATCCATGTTGGCC 720  
 QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
 Db 721 TCTCGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTGATCTT 780  
 QY 257 LeuArgLeuAsp 261  
 Db 781 CTACGTCGCTCGAT 795

## RESULT 8

US-10-267-763-3  
 ; Sequence 3, Application US/10267763  
 ; Publication No. US20030077687A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gorlach, Jorn  
 ; APPLICANT: Boyes, Douglas  
 ; APPLICANT: Davis, Keith  
 ; APPLICANT: Hamilton, Carol  
 ; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Kloti, Andreas  
 ; APPLICANT: Woessner, Jeffrey  
 ; APPLICANT: Zayed, Adel  
 ; APPLICANT: Ascenzi, Robert  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
 ; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE  
 ; FILE REFERENCE: 2035DIV1  
 ; CURRENT APPLICATION NUMBER: US/10/267,763  
 ; CURRENT FILING DATE: 2002-10-09  
 ; PRIOR APPLICATION NUMBER: 09/610,040  
 ; PRIOR FILING DATE: 2000-07-05  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 1006  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis  
 US-10-267-763-3

Alignment Scores:  
 Pred. No.: 3,07e-96 Length: 1006  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservative: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 15 Gaps: 2

US-10-624-061-16 (1-261) x US-10-267-763-3 (1-1006)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSergly-----AsnValTyrThrLeuAla 18  
 Db 64 ATGGCAAGATCTTCGAATCGATTCGGTTCGTTGTTCCCAATGACTGAGTCTTGAC 123  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArg 38  
 Db 124 TTAATCAGAGAAATCGTTGATGAGCAAGACACCATCTCTTCAGCTTGTATCGAGAGA 183  
 QY 39 AlalysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 Db 184 GCTAAGTTTCCACTCAATCTCTCGCTTCGAGGAATCTGTTGCTAGATTCGGAAGT 243  
 QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 Db 244 TCTCTCTCTCACTGAGTTTTCGTCAGAGACAGACAAATCATCAAGTAAGTAGGA 303  
 QY 79 ArgTyrLysAsnProGluGluAsnAlaPheProGluAsnLeuProProSerIleVal 98  
 Db 304 AGATATCAATACCGGAAGAGATCTCTTCTCTTGAGAACATCTCTCACTCGGTTTTT 363  
 QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLeuLys 118

Db 364 CCTACGCACAAATATCCATCGGCTTTGCACCCCTAAGGCTCTATCTGTTAACTTAACAAA 423  
 QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138  
 Db 424 CAATCTGGGATATTTACTTTTAAAGAAATTCCTTCTTTGTTTGTCAACCTGGCGATGAT 483  
 QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158  
 Db 484 GGCAACTATCCATCAACTGCTGTAGTGTCTGCTGCTGTCTGCTGTCTTACAGCTCTTTCGAGAAGG 543  
 QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 Db 544 ATTCACTACGGTAAATTTGTAGCTGAGGTCAATTCAGAGATGCTCCCAAGATACGAG 603  
 QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198  
 Db 604 CTTGCGATTCGCGCTCAGGATAGAGAGGCTTTGATGAAGCTGTGACGTTTGAGAAAGTA 663  
 QY 199 GluGluThrValArgLysArgValGluLysLysAlaValPheGlyGlnGluValAsn 218  
 Db 664 GAAGAAATGTTTAAAGAGAGAGTGCAGAAAGACAGAAACCGTTTGGACAAGAGTAAAA 723  
 QY 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236  
 Db 724 TTCAACTCTGGCTATGGCGATGAGAGTAAAGAGAAGTATAAAGTGATCCATGTTGCC 783  
 QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
 Db 784 TCTCGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTGATCTT 843  
 QY 257 LeuArgLeuAsp 261  
 Db 844 CTACGTCGCTCGAT 858

## RESULT 9

US-10-267-763-9/c  
 ; Sequence 9, Application US/10267763  
 ; Publication No. US20030077687A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gorlach, Jorn  
 ; APPLICANT: Boyes, Douglas  
 ; APPLICANT: Davis, Keith  
 ; APPLICANT: Hamilton, Carol  
 ; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Kloti, Andreas  
 ; APPLICANT: Woessner, Jeffrey  
 ; APPLICANT: Zayed, Adel  
 ; APPLICANT: Ascenzi, Robert  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
 ; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE  
 ; FILE REFERENCE: 2035DIV1  
 ; CURRENT APPLICATION NUMBER: US/10/267,763  
 ; CURRENT FILING DATE: 2002-10-09  
 ; PRIOR APPLICATION NUMBER: 09/610,040  
 ; PRIOR FILING DATE: 2000-07-05  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 9  
 ; LENGTH: 1006  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-10-267-763-9

## Alignment Scores:

Pred. No.: 3,07e-96 Length: 1006  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservative: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 15 Gaps: 2

```
US-10-624-061-16 (1-261) x US-10-267-763-9 (1-1006)
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (432)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (514)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (516)
US-09-454-279-5
Alignment Scores:
Pred. No.: 4,31e-90
Score: 767.50
Percent Similarity: 95.29%
Best Local Similarity: 95.29%
Query Match: 57.32%
DB: 9
Gaps: 0
US-10-624-061-16 (1-261) x US-09-454-279-5 (1-525)
QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20
Db 16 ATGCCAAACGACGACGACAAAGTCTTGATCTGGGATGTGTACACGCTAGCTTCTGTG 75
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40
Db 76 AGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTTATGGTCTCATTTGAGAGAGCCAA 135
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60
Db 136 TTCCTTAGCAATTTCTCACCTATGATGAAAGATGTCTCAATCCAGGGTTTGTGGC 195
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80
Db 196 TCATTTGGTGAATTTGTTTAAAGATACAGAGCCATTCAGGTAAGGCTGGAAGATAC 255
QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100
Db 256 AAAAACCCCTGAAGAAAACGCCCTTCTCCAGAAAATTTACCACCATCAATTTGCCATCT 315
QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerile 120
Db 316 TACTCTTCAACACAGTTTTCATCTCCTGCTGCTCTTCAATTAACATAACAAGTCATCT 375
QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140
Db 376 GGGAAAATGTATTTCAAGAGAGTTACTTCCATTGCTTGTCTTCTTCCGGGTGATGGNAAC 435
QY 141 TyrAlaGln-ThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgGileHi 160
Db 436 TATGCGCAAACTGCGCTAATGACCT-TCATTATTGCGATC-ATCTTAGAGGAT-CA 492
QY 160 sTyrGlyLysPheValAlaGluValLys 169
Db 493 CTATGGAAGTTTGTAGCTGANGGAAA 520
RESULT 11
US-10-624-061-5
; Sequence 5, Application US/10624061
; Publication No. US20040019929A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Farnodu, Omelayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/10/624,061
; NUMBER OF SEQ ID NOS: 22
```

; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: US/09/454,279  
; PRIOR FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 5  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (432)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (514)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (516)  
US-10-624-061-5

Alignment Scores:  
Pred. No.: 4,31e-90 Length: 525  
Score: 767.50 Matches: 162  
Percent Similarity: 95.2% Conservative: 0  
Best Local Similarity: 95.2% Mismatches: 7  
Query Match: 57.32% Indels: 4  
DB: 16 Gaps: 0  
US-10-624-061-16 (1-261) x US-10-624-061-5 (1-525)

QY 1 MetAlaLysAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
DB 16 ATGGCCAAAGCAGCAGACAAAGTCTGATTCGGGAATGTACACGGTAGCTTCGTG 75  
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLys 40  
DB 76 AGAGAGGATTGGTTAGCAAGAGGATACCATCATTTATGCTCTCATTTAGAGAGCCAG 135  
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
DB 136 TTCCTAGCAATTCACACCATATGATGAAAGATGCTCAAAATCCAGGGTTTTGTGGC 195  
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
DB 196 TCATTGGTGAATTTGTTTAAAGATACAGAGCCATTCAAGTAGGCTGGAGATAC 255  
QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProSerIleValProSer 100  
DB 256 AAAAACCCCTGAAGAAAGCCCTCTCTCCAGAAAATTTACCACCATCAATTGTGCCATCT 315  
QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLysSerIle 120  
DB 316 TACTCTTCAACAGTTTTCATCTCTGGTCTGCTCTCAATTACATTAACAGTCACT 375  
QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140  
DB 376 GGGAAAATGTATTTCAAGAGATTACTTCCATTGCTTCTACTTCGGGTGATGAGNAAC 435  
QY 141 TyrAlaGln-ThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgIleHi 160  
DB 436 TATGGCAAACTCAGCTAATGACCT-TCATTATTGAGTTC-ATCTCTAGAAGGAT-CA 492  
QY 160 sTyrGlyLysPheValAlaGluValLys 169  
DB 493 CTATGGAAGTTTGTAGCTGANGNGAAA 520

## RESULT 12

US-09-938-842A-737  
; Sequence 737, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 737  
; LENGTH: 1005  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-737

Alignment Scores:  
Pred. No.: 1,37e-80 Length: 1005  
Score: 698.00 Matches: 131  
Percent Similarity: 68.08% Conservative: 46  
Best Local Similarity: 50.38% Mismatches: 75  
Query Match: 52.13% Indels: 8  
DB: 9 Gaps: 1  
US-10-624-061-16 (1-261) x US-09-938-842A-737 (1-1005)

QY 10 AppSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29  
DB 223 GATGAGAGTGAGAGTTTGACTCTCTGAAGGTATTAGAAACTCTTTGATCCGTCAGAGGAC 282  
QY 30 ThrIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49  
DB 283 AGCATTAATTTGGGCTATTGGAGAGCCAGTACTGTTCATCATCTGATACATTATGAT 342  
QY 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69  
DB 343 CCTACTGCTTTTGACATGATGCTTCAATGGTTCTTTGGTTGAGTACATGTTAAAGGC 402  
QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89  
DB 403 ATGAGAGCTTCACGCTTAAGTTGTTAGGTCTCTGATGATGAACATCTTTCTTC 462  
QY 90 ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
DB 463 CCTGATGATCTACAGAGCTATGTGCTCTCTCTTCAAGTACCCAAAGGTGTTCATTTT 522  
QY 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129  
DB 523 GCTGCTGATTCGATTAACATTAACAAAGAGATATGGAACATGTACTTCAGAGACCTTGT 582  
QY 130 ProLeuLeuAlaThrSerGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
DB 583 CCAAGACTTGAAGAAAGCGGATGATGTTAATCTACGGCTCAACAGCTGTCTGTGACGCT 642  
QY 150 SerLeuLeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLys 169  
DB 643 APTCCCTTCAGTGTCTCTCAAGAGAAATCCATACCGTAAATTTGTTGCAAGAGCTAAA 702  
QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189  
DB 703 TTTCAGGCTCACCGAAGCATACGAGTCCGCCATCAAGCAAGATAGGATGCACTG 762  
QY 190 MetLysLeuLeuThrPheThrSerValGluThrValArgLysArgValGluLysLys 209  
DB 763 ATGGATATGCTGACATTCCTGAGATGCGAATCGGATAAAGAGAGAGATTGAGTGA 822  
QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAsnAsnGluAsnArg 229

```

Db      823 ACCGGAACATACCGGCAAGAGTGAAGTGGATGGATGGAGGAGAAAGAAAGAGAA 882
Qy      230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
Db      883 GAAGGAATGAATCATCTTTACAAATCATCGATCTTACGTTAGTTGGTACTTATATGGA 942
Qy      242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuValArgLeuAsp 261
Db      943 GATTGGATCATGCCCTTTACAAAGAGGTTCAAGTGGAGTACTTCTGTCAGAGACTGGAC 1002

RESULT 13
US-09-938-842A-737
; Sequence 737, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreppe, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; SAME, AND METHODS OF USE
; FILE REFERENCE: SCRI1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 737
; LENGTH: 1005
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-737

Alignment Scores:
Pred. No.: 1,37e-80 Length: 1005
Score: 698.00 Matches: 131
Percent Similarity: 68.08% Conservative: 46
Best Local Similarity: 50.38% Mismatches: 75
Query Match: 52.13% Indels: 8
DB: 11 Gaps: 1

US-10-624-061-16 (1-261) x US-09-938-842A-737 (1-1005)
Qy      10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
Db      223 GATGAGAGTGAGAGTTTGACTCTTGAAGGTATTAGAACTCTTTGATCGGTCAAGAGGAC 282
Qy      30 ThrIleIleTyrGlyLeuLeuGluArgAlaLysPhePheProSerAsnSerHisThrTyrAsp 49
Db      283 AGCATATTATTTGGGTATTGGAGAGAGCCAGTACTGTGTACAACTGCTGATCTATTGAT 342
Qy      50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValLysAsn 69
Db      343 CCTACTGCTTTTCACATGGATGTTCTCAATGGTTCTTTGGTTCAGTACATGGTTAAAGGC 402
Qy      70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89
Db      403 ACTGAGAAGCTTCACGTCAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 462
Qy      90 ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
Db      463 CCTGATGATCTACAGAGCCTATTGTCCTCTCTTCAGTACCAAGGTGTGCAATTT 522
Qy      110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129
Db      523 GCTGCTGATTCGATAACATAAACAAGAGATATGGAACATGATGATCTTCAGAGACCTTGT 582
Qy      130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149

```

```

Db      583 CCAAGACTTTGTGAAGAAAGCGGATGATGGTAATACCGCTCAACAGCTGTCTGTGAGCT 642
Qy      150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys 169
Db      643 ATTCGCTTCAGTGTCTCTCAAAGAGAAATCCATTACGTTAAATTTGTTGCAAGAGCTAAA 702
Qy      170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189
Db      703 TTTCAAGCCCTCACCCGGAAGCATACGAGTCGCGCATCAAGCAAGATAGGATGCACTG 762
Qy      190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
Db      763 ATGATATGTCGATCTCCGACTGTGGAAGATGCGATAAAGAGAGAGTTGAGATGAAA 822
Qy      210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAspAsnAspAsnGluAsnArg 229
Db      823 ACCCGAACATACGGGCAAGAGTGAAGTTGGGATGGAGGAGAAAGAAAGAAAGAA 882
Qy      230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
Db      883 GAAGGAATGAATCTCATGTTTACAAATCATGCTCCGATCTTAGTTGCTGACTTATATGGA 942
Qy      242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeuAsp 261
Db      943 GATTGGATCATGCCCTTTACAAAGAGGTTCAAGTGGAGTACTTGTCTCAGAGACTGGAC 1002

RESULT 14
US-10-267-763-2
; Sequence 2, Application US/10267763
; Publication No. US200300077687A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1207
; TYPE: DNA
; ORGANISM: Arabidopsis
US-10-267-763-2

Alignment Scores:
Pred. No.: 3.32e-80 Length: 1207
Score: 696.00 Matches: 131
Percent Similarity: 68.08% Conservative: 46
Best Local Similarity: 50.38% Mismatches: 75
Query Match: 51.99% Indels: 8
DB: 15 Gaps: 1

US-10-624-061-16 (1-261) x US-10-267-763-2 (1-1207)
Qy      10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
Db      235 GATGAGAGTGAGAGTTTGCATCTCTTGAAGGTATTGAACTCTTTGATCCGTCAGAGGAC 294
Qy      30 ThrIleIleTyrGlyLeuLeuGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49

```

```

Db      295 AGCAATTATATTTGGGCTATTGGGAGAGAGCCAAAGTACTGTGTACAAATGCTGTAATCTTATGAT 354
QY      50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69
Db      355 CCTACTGCTTTTGCACATGATGGTTTCAATGCTTCTTTGGTTGAGTACATGCTTAAAGGC 414
QY      70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89
Db      415 ACTGGAAGCTTTCACGCTTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 474
QY      90 ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
Db      475 CCTGATGATCTACAGAGCCTATGTTGCTCTCTCTTCAAGTACCCAAAGGTGTGCAATTT 534
QY      110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129
Db      535 GCTGCTGATTCGATAAACAATAACAAGAGATATGGAACATGTACTTTCAGAGACCTTGT 594
QY      130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149
Db      595 CCAAGACTTGTGAAGAGCGCATGTGTAAGTAACTACGCTCAACAGCTGTCTGTGACGCT 654
QY      150 SerLeuLeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLys 169
Db      655 ATCTGCTTTCAGTCTCTCAAGAGAAATCCATTACCGTAAATTTGTTGCAAGAGCTAAA 714
QY      170 PheArgAspAlaProGluAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189
Db      715 TTTCAGGCTTACCCGAGCATACGAGTCGCGCATCAAGAGCACAAGATAAGGATCGACTG 774
QY      190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
Db      775 ATGGATATGCTGACATTCGCACTCTGGAAGATCGATAAAGAGAGAGATTGAGATGAA 834
QY      210 AlaValValPheGluGluValAsnLeuAsnSerAspAsnAspAsnAspAsnGluAsnArg 229
Db      835 ACCCGAACATACGGCAAGAGTGAAGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 894
QY      230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
Db      895 GAAGGGAATGATCTCATGTTTACAAATCATGTCGATCTTAGTTGTTGATATATGGA 954
QY      242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeuAsp 261
Db      955 GATTGCTATGCTCTTTAAACAAAGAGGTTCAAGTGGAGTACTTGGCTCAGAAGACTGGAC 1014

RESULT 15
US-10-267-763-8/c
; Sequence 8, Application US/10267763
; Publication No. US20030077687A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Klotz, Andreas
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1207
; TYPE: DNA

```

```

; ORGANISM: Arabidopsis thaliana
US-10-267-763-8

```

## Alignment Scores:

```

Pred. No.: 3 32e-80 Length: 1207
Score: 696.00 Matches: 131
Percent Similarity: 68.08% Conservative: 46
Best Local Similarity: 50.38% Mismatches: 75
Query Match: 51.98% Indels: 8
DB: 15 Gaps: 1

```

```
US-10-624-061-16 (1-261) x US-10-267-763-8 (1-1207)
```

```

QY      10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
Db      973 GATGAGAGTGAGAGTTTGAATCTCTTGAAGGTATTAGAAACTCTTTGATCCGTCAGAGGAC 914
QY      30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49
Db      913 AGCATTTATATTGGGCTATTGGAGAGAGCCAAAGTACTGTTACATGCTCATCTTATGAT 854
QY      50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69
Db      853 CCTACTGCTTTTGACATGATGTTTCAATGCTTCTTTGGTTGAGTACATGTTTAAAGGC 794
QY      70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89
Db      793 ACTGGAAGCTTACACCTAAGGTTGTAGTTTAAAGTCTCTGATGAACATCTTTCTTTC 734
QY      90 ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
Db      733 CCTGATGATCTACAGAGCCTATGTTGCTCTCTCTTCAGTACCCAAAGGTGTGTCATTT 674
QY      110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129
Db      673 GTGCTGATTCGATAAACAATAACAAGAGATATGGAACATGTACTTTCAGAGACCTTGT 614
QY      130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149
Db      613 CCAAGACTTGTGAAGAAAGCGCATGTGTTAACTACGGCTCAACAGCTGTCTGTGACGCT 554
QY      150 SerLeuLeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLys 169
Db      553 ATCTGCTTTCAGTCTCTCTCAAGAGAAATCCATTACGGTAAATTTTTCGCAAGACTAAA 494
QY      170 PheArgAspAlaProGluAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189
Db      493 TTTCAGGCTTACCCGAGCATACGAGTCCGCCCATCAAGGACACAGATPAAGATCGACTG 434
QY      190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
Db      433 ATGGATATGCTGACATTCGCACTCTGGAAGATGCGATAAAGAGAGAGATTGAGATGAAA 374
QY      210 AlaValValPheGluGlnGluValAsnLeuAsnSerAspAsnAspAsnAspAsnGluAsnArg 229
Db      373 ACCCGAACAATACGGCAAGAGTGAAGTGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAA 314
QY      230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
Db      313 GAAGGGAATGATCTCATGTTTACAAATCATGTCGATCTTAGTTGGTGGTACTTATATGGA 254
QY      242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeuAsp 261
Db      253 GATTGCTATGCTCTTTAAACAAAGAGGTTCAAGTGGAGTACTTGGCTCAGAAGACTGGAC 194

```

Search completed: September 25, 2004, 06:20:33

Job time : 520 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 25, 2004, 02:58:27 ; Search time 2777 Seconds  
(without alignments)  
2806.636 Million cell updates/sec

Title: US-10-624-061-16  
Perfect score: 1339  
Sequence: 1 MAXAAEQPSDGNVTVLASV.....NWVPLTKEVQVYLLRLD 261

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_p2n.model -DRV=xlh  
-Q=/cgm2\_1/US10624061/runat\_24092004\_092407\_16924/app\_query.fasta\_1.455  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORW=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10624061 @CGN 1.1 3437 @runat\_24092004\_092407\_16924 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:  
2: em\_esthum:  
3: em\_estin:  
4: em\_estnu:  
5: em\_estov:  
6: em\_estpl:  
7: em\_estro:  
8: em\_hic:  
9: gb\_est1:  
10: gb\_est2:  
11: gb\_hic:  
12: gb\_est3:  
13: gb\_est4:  
14: gb\_est5:  
15: em\_estfun:  
16: em\_estom:  
17: em\_gss\_hum:  
18: em\_gss\_inv:  
19: em\_gss\_pln:  
20: em\_gss\_vrt:  
21: em\_gss\_fun:  
22: em\_gss\_mam:  
23: em\_gss\_mus:  
24: em\_gss\_pro:  
25: em\_gss\_rod:  
26: em\_gss\_phg:  
27: em\_gss\_vrl:  
28: gb\_gss1:

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	917	68.5	563	10	AW666427 sk36d12.y
2	895	66.8	648	10	AW774625
3	876	65.4	512	12	BG363365 sac13f08.
4	831	62.1	654	10	AW775842
5	831	62.1	696	12	B1273174
6	794	59.3	605	10	BF649665 NF081B12F
7	755	56.4	620	10	BF009709
8	723.5	54.0	605	12	B1420527
9	721	53.8	477	9	A1973566
10	721	53.8	560	10	BE821647
11	717	53.5	529	14	CB827129
12	702	52.4	422	12	BG239446
13	679.5	50.7	810	14	CD427426
14	672	50.2	560	12	BM033460
15	671	50.1	1196	11	AY107907
16	670.5	50.1	1222	11	AY103806
17	670	50.0	748	14	CD824727
18	667	49.8	726	14	CA800612
19	661.5	49.4	710	14	CF475712
20	649.5	48.5	792	12	BG309831
21	622.5	46.5	660	13	EX233798
22	618.5	46.2	821	14	CK196770
23	609.5	45.5	867	14	CB619882
24	608.5	45.4	738	13	BQ989273
25	604.5	45.1	524	14	CB828627
26	590.5	44.1	531	14	CB828445
27	590.5	44.1	557	14	CB829043
28	586.5	43.8	516	14	CB827444
29	583.5	43.6	680	9	AJ612241
30	580	43.3	885	13	BU693059
31	579.5	43.3	761	14	CF480460
32	579	43.2	639	10	BE660542
33	578.5	43.2	501	14	CB828062
34	573.5	42.8	736	14	CF443707
35	573	42.8	698	10	BE577489
36	572.5	42.8	549	14	CB828614
37	571.5	42.7	529	14	CB826768
38	571	42.6	680	13	CA080410
39	567.5	42.4	541	14	CB829238
40	564	42.1	683	14	CF481222
41	563.5	42.1	725	14	CF486896
42	561	41.9	696	14	CD919148
43	559	41.7	595	10	BE445201
44	552.5	41.3	830	14	CB622596
45	550.5	41.1	711	10	BF430733

ALIGNMENTS

RESULT 1  
AW666427  
LOCUS  
DEFINITION  
AW666427 563 bp mRNA linear EST 03-DEC-2001  
sk36d12.y1 Gm-cl028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-cl028-4272 5', similar to SM:CHMU\_ARATH P42738 CHORISMATE MUTASE  
PRECUSOR ; mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AW666427 GI:7478842  
Glycine max (soybean)  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

## REFERENCE

## AUTHORS

1 (bases 1 to 563)

Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,  
Khanna, A., Bolla, B., Marra, M., Hillier, L., Aucuba, T., Martin, J.,  
Beck, C., Wylie, J., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
Bowers, J., Person, B., Swallow, T., Gibbons, M., Pape, D., Harvey, N.,  
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)

## TITLE

## JOURNAL

## COMMENT

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
High quality sequence stop: 418.

## FEATURES

## source

1..563

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1028-4272"

/tissue\_type="roots of 'Supernod' plants"

/lab\_host="DH10B"

/clone\_lib="Gm-c1028"

/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:  
XhoI. The mRNA was isolated from roots of Glycine max  
'Supernod' plants generously donated by Dr. Gary Stacey.  
The seedlings were inoculated with Bradyrhizobium  
japonicum, strain USDA110 prior to harvest. StrataGene's  
cDNA synthesis kit (catalog number 200401) was used to  
synthesize the cDNA. First-strand synthesis was performed  
with 5-methyl dCTP, hence the ligated cDNA was  
hemimethylated. A modification of StrataGene's  
first-strand synthesis primer was used. An 'anchor'  
nucleotide (V=A,C, or G) was added to the 3' end of the  
primer [GAGAGAGAGAGAGAGAGAGAGAGAG(T)18V] to anchor  
the primer at the 5' end of the poly(A) tract. After  
second-strand synthesis, the cDNA ends were filled in with  
cloned Pfu DNA polymerase, ligated to EcoRI adapters and  
subsequently phosphorylated. The XhoI site within the  
first-strand synthesis primer was then restricted by  
digestion with XhoI; all XhoI sites in the cDNA would be  
protected by their hemimethylated status. The cDNA  
constructs were size-fractionated with a 500bp cutoff,  
using GibcoBRL Life Technologies' cDNA Size Fractionation  
column. The column eluent was then ligated into  
StrataGene's pBluescript II XR Predigested vector  
(pBluescript II SK(+)) that has been digested with EcoRI  
and XhoI, and phosphorylated by StrataGene. Both the  
white and blue colonies appear to contain recombinant  
plasmids with cDNA inserts, based on size (n=25). This  
library was constructed by Dr. Paul Keim and Dr. Virginia  
Coryell."

## ORIGIN

## Alignment Scores:

Pred. No.:	1.16e-103	Length:	563
Score:	917.00	Matches:	176
Percent Similarity:	97.85%	Conservative:	6
Best Local Similarity:	94.62%	Mismatches:	4
Query Match:	68.48%	Indels:	0
DB:	10	Gaps:	0

US-10-624-061-16 (1-261) x AW666427 (1-563)

Qy

25 ValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLysPheProSerAsn 44

Db	6	GTTAGGCAAGAGATACCATCATTTATGGTCTCATTTGAGAGAGCCAAAGTTCCTAGCAAT	65
Qy	45	SerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGlu	64
Db	66	CCTCACACCATCATGATGAAGAGATGCTCAATCCAGGGTTTTTGTGGCTCATTTGGTAGAA	125
Qy	65	PheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyValGlyTyrLysAsnProGlu	84
Db	126	TTTGTGTTAGATACAGAGCCATTCAGCTGAGGCTGGAGATACAAAACCCCTGAA	185
Qy	85	GluAsnAlaPhePheProGluAsnLeuProProSerIleValProSerTyrSerPheLys	104
Db	186	GAAGAGCCCTTCTCCAGAAAAATTTACCACCATCAATTTGTGCATCTTACTCCITCAA	245
Qy	105	GlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyr	124
Db	246	CAGTTTTCATCTCGAGCTGCTTCAATTACATTAACAAGTCCATCTGGAAATGTAT	305
Qy	125	PheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThr	144
Db	306	TTCCAAAGAGTTACTTCCATTGCTTGTCTACTTCGGGGGATGATGGAACATATGCACAACT	365
Qy	145	AlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPhe	164
Db	366	GCAGCTAAATGATCTTTTATTTGAGGCGCATCTCTAGAAGGANTTCATATGGAAGTTT	425
Qy	165	ValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLys	184
Db	426	GTAGCTGAGGTGAATTCANGATGCTCTCTCAAGACTACGAGCTTTTATTCGAGCTAAG	485
Qy	185	AspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGluThrValArgLys	204
Db	486	GATAAAGAGAGATTGATGAATTTGGTGATCATTTACAGCGTTTGAAGAGACAGTGGAGAA	545
Qy	205	ArgValGluLysLysAla 210	
Db	546	GAGGTGTAAGAAGAGGCC 563	

## RESULT 2

## AW774625

## LOCUS

## DEFINITION

BST333776 KV3 Medicago truncatula cDNA clone pkV3-23M13, mRNA

sequence.

AW774625

AW774625.1

GI:7718542

EST.

Medicago truncatula (barrel medic)

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

1 (bases 1 to 648)

Vandenbosch, K., Hurt, J., Moore, J., Beremand, P., Peng, H., Ellis, L.,

Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and

Fraser, C.M.

ESTs from roots of Medicago truncatula after Rhizobium inoculation

Unpublished (1999)

Contact: Vandenbosch K

Department of Plant Biology

University of Minnesota

220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA

Tel: 612 624 2755

Fax: 612 625 1738

Email: kvandenbosch@umn.edu

Texas A&M EST name: T258044c

TIGR sequence name: MTEBC79TK

More information is available at:

http://chrysis.tamu.edu/medicago

Seq primer: SKmod (CTA GAA CTA gfg gAT CC).

Location/Qualifiers

1..648

## FEATURES

## source

```

/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pKV3-23M13"
/tissue_type="Seedling roots"
/dev_stage="3 days post-inoculation with Sinorhizobium
melliotti"
/lab_host="E. coli strain XLOLR"
/clone_lib="KV3"
/notes="Vector: pBluescript SK -; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unisap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XLOLR cells."

ORIGIN
Alignment Scores:
Pred. No.: 8,02e-101 Length: 648
Score: 895.00 Matches: 167
Percent Similarity: 87.50% Conservative: 22
Best Local Similarity: 77.31% Mismatches: 27
Query Match: 66.84% Indels: 0
DB: 10 Gaps: 0

US-10-624-061-16 (1-261) x AW774625 (1-648)
QY 34 GlyLeuIleGluAlaGlySerProSerAsnSerHisThrTyrAspGluLysTyrAla 53
DB 1 GGCTTATTGAGAGATCAAAAGTTCCCTTTAATCTCAAACTATGATCAGAAATTATCT 60
QY 54 GlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsnThrGluAlaIle 73
DB 61 CAAATCCCTGGTTTGTGGCTCATTTGGTGGTTGTTTTCACCAATCTAGATGTT 120
QY 74 GlnAlaLysAlaGlyArgTyrLysAsnProGluGluAlaPhePheProGluAsnLeu 93
DB 121 CAGCTAAGCTGGAAGATACAAAACCTCGAGAGATCCCTCTTCCAGAAATTTA 180
QY 94 ProProSerIleValProSerTyrSerPheGlyGlnPheLeuHisProGlyAlaAlaSer 113
DB 181 TCATTGTCACTTGTTCCTATCCCTTCACAAAGTTTTTGATCCTCTGGAGCGCTCG 240
QY 114 IleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeuProLeuAla 133
DB 241 ATTAACATAAACAGTCTTATGAAATGATCTTTCATGACCTGCTTCCATTGTCGTT 300
QY 134 ThrSerGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGln 153
DB 301 GCCTCCGGTGATGATGCAACTATGCAAACTGCAGTAGCAGCTCTCATTTGTCAG 360
QY 154 SerIleSerArgGlnIleHisTyrGlyLysPheValAlaGluValLysPheArgAspAla 173
DB 361 GCCATTTCTAAAGAGTTCATTATGGAAGATTTGTAGCTGAAGTGAATTCAGGAATCT 420
QY 174 ProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeu 193
DB 421 CCTCAAGACTATGAGCTTTAATTCGCTCAGGACAAAGAGGTTGATGAATTTGTTG 480
QY 194 ThrPheThrSerValGluGluThrValArgLysArgValGluLysLysAlaValAlaPhe 213
DB 481 ACATTTAAGAGTGTGAAGAGATGTTGAAGAAAAGGGTTGAAAAGAGGCAACATGTTT 540
QY 214 GlyGlnGluValAsnLeuAsnSerAspAsnAspAsnGluAsnArgLysPheAspPro 233
DB 541 GGGCAGGATATPACCTTTACAGCAGCTGATGACGTAAAGGTAAACAAAGTTTGTCCA 600
QY 234 SerValAlaSerSerLeuTyrLysAsnTrpValIleProLeuThrLys 249
DB 601 TCAGTGGCTTCTAAGTGTATGAGAAATGGGTAAATACCTCTTACTAAG 648

```

```

RESULT 3
BG363365
LOCUS
DEFINITION
sac19f08.y1 Gm-cl051 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl051-2943 5' similar to TR:Q9S7H4 Q9S7H4 CHORISMATE MUTASE
; mRNA sequence.
ACCESSION
BG363365 GI:13252454
VERSION
BG363365.1
SOURCE
Glycine max (soybean)
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
1 (bases 1 to 512)
Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,L., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 397.
Location/Qualifiers
1..512
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-2943"
/tissue_type="floral meristematic mRNA"
/lab_host="DH10B"
/clone_lib="Gm-cl051"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from floral
meristematic mRNA provided by Dr. Halina Knap of Clemson
University. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1,32e-98 Length: 512
Score: 876.00 Matches: 169
Percent Similarity: 99.41% Conservative: 0
Best Local Similarity: 99.41% Mismatches: 1
Query Match: 65.42% Indels: 0
DB: 12 Gaps: 0

US-10-624-061-16 (1-261) x BG363365 (1-512)
QY 7 GlnSerProAspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArg 26
DB 3 CAAAGTCCTGATCTGGGAATGTGACACGTAGCTTCTCTGAGAGAGATTGTTAGG 62
QY 27 GlnGluAspThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHis 46

```

```

Db 63 CAAGAGGATACCATATTTATGGTCTCATTTAGAGAGCCAAAGTTCCTAGCAATTTCTCAC 122
QY 47 ThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheVal 66
Db 123 ACCTATGATGAAGAAGTATGCTCAATCAACAGGGTTTTTGTGGCTCATTTGGTGAATTTGTT 182
QY 67 ValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsn 86
Db 183 GTTAAGAATACAGAGGCCATTCAGCTAGGCTGGAAGATACAAAGAACCTGAAGAAAC 242
QY 87 AlaPhePheProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPhe 106
Db 243 GCTTCTCCAGAAATTTACACATCAATGTGGCCATCTTACTCTTCAACAGTTT 302
QY 107 LeuHisProGlyAlaAlaSerIleAsnLysSerIleTyrLysMetTyrPheLys 126
Db 303 TTGCATCTCTGGTCTGCTTCAATTAACATAAACAAGTCCATCTGGAAATGTATTTCAA 362
QY 127 GluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAla 146
Db 363 GAGTTACTTCTCAATGCTTCTACTTCTGGGTGATGATGCAACTATGCCAAACATGCAGCT 422
QY 147 AsnAspLeuSerLeuLeuGlnSerIleSerArgGlyIleHisTyrTyrGlyLysPheValAla 166
Db 423 AATGACCTTCTTCAATTTATGTCAGTCCATCTCTAGAGGATTCATATGGAAGTTTGTAGCT 482
QY 167 GluValLysPheArgAspAlaProGlnAsp 176
Db 483 GAGTGAATTCAGGAGTCTCTCTCAAGAC 512

```

## RESULT 4

AW775842  
LOCUS  
DEFINITION  
EST3343907 DSIL Medicago truncatula cDNA clone pDSIL-3E8, mRNA

ACCESSION  
AW775842  
VERSION  
AW775842.1 GI:7765655

KEYWORDS  
SOURCE  
ORGANISM  
Medicago truncatula (barrel medic)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.

REFERENCE  
1 (bases 1 to 654)  
Autors  
Pedorova, M., Pierson, B.L., Samac, D.A., Vance, C.P., Gantt, G.S.,  
Peng, H., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B.,  
Hansen, I.S., Holt, I.B. and Fraser, C.M.  
ESTs from leaves of Medicago truncatula after inoculation with  
Colletotrichum trifolii

JOURNAL  
COMMENT  
Unpublished (2000)  
Contact: Deborah A. Samac  
Department of Plant Pathology  
University of Minnesota  
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA  
Tel: 612 625 1243  
Fax: 651 649 5058  
Email: debbys@puccini.crl.umn.edu  
Minnesota sequence name: M259175e  
TIGR sequence name: MTF28TK  
More information is available at:  
http://chrysis.tamu.edu/medicago  
Seq primer: SKmod (CTA GAA CTA Gtg GAT CC).

## FEATURES

source  
1..654  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/cultivar="genotype A17"  
/db\_xref="taxon:3880"  
/clone="pDSIL-3E8"  
/tissue\_type="leaves infected with Colletotrichum trifolii"

/dev\_stage="cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii"  
/lab\_host="E. coli strain XL0LR"  
/clone\_lib="DSIL"  
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA from cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii. The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in XL0LR cells. Note: EST may be of fungal origin."

## ORIGIN

Alignment Scores:  
Pred. No.: 7,92e-93 Length: 654  
Score: 831.00 Matches: 157  
Percent Similarity: 88.83% Conservative: 18  
Best Local Similarity: 79.70% Mismatches: 22  
Query Match: 62.06% Indels: 0  
DB: 10 Gaps: 0  
US-10-624-061-16 (1-261) x AW775842 (1-654)

QY 13 AsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGluAspThrIle 32  
Db 64 AATGAATATATCTCTTGATACAGTGGGAGACTTTGGTTAGACAAGATACAAATGTT 123  
QY 33 TyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAspGluLysTyr 52  
Db 124 TTGTGCTTATTGAGAGATCAAGATTTCTCTTTAATTTCTCAAACTTATGATCAGAATTAT 183  
QY 53 AlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsnThrGluAla 72  
Db 184 CTTCAAAATCCCTGGTTTTTGTGGCTCATTTGTGATTTGTTTTCACCAATCTCAGATT 243  
QY 73 IleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsn 92  
Db 244 GTTCAAGCTAAGCTGGAAGATACAAAACCTCGAAGATCCCTCTTCCAGAAAT 303  
QY 93 LeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAla 112  
Db 304 TTATCATTTGCTCATTTGTTCCATCTTACCCCTTACAAAAGTTTTTGCATCTCGAGCCGCT 363  
QY 113 SerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeu 132  
Db 364 TCGATTACATTAACCAAGTCTTATGGAATAAGTACTTTTCATGACTGCTTCCATTTGTC 423  
QY 133 AlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeu 152  
Db 424 GTTGCTTCCGGTGATGTCGCAACTATGCACAACTGCGAGCTAGCGACTCTCTCATTTGT 483  
QY 153 GlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLysPheArgAsp 172  
Db 484 CAGGCCATTTCTAAAGAGATTCATTATGGAAGTTTGTAGCTGAAGTGAATTCAGGAA 543  
QY 173 AlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeuMethLysLeu 192  
Db 544 TCTCCTCAAGACTATGAGCTTTAATTCGGTCTAAGGACAAAGAGGGTTGATGAATTTG 603  
QY 193 LeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209  
Db 604 TTGACATTTAAGAGTGTGTAAGAGATGTTAAAGAAAGGGTTGAAAGAAAG 654

## RESULT 5

BI273174  
LOCUS  
DEFINITION  
NF091B12FL1F1096 Developing flower Medicago truncatula cDNA clone  
BI273174  
ACCESSION  
BI273174  
VERSION  
BI273174.1 GI:14883109

696 bp mRNA linear EST 18-JUL-2001

```

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 696)
Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula flower library
Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert length: 696 Std Error: 0.00
Plate: 091 row: B column: 12
Seq primer: TCACACAGGAACAGCTATGAC.

FEATURES
source
1..696
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF091B12FL"
/tissue_type="Developing flowers"
/dev_stage="Developmentally pooled. Contains a mixture of
very young, developing, fully-opened flowers and
in early transition into pods."
/clone_lib="Developing flower"
/notes="Vector: lambda Zap; cDNA was prepared from polyA+
enriched, pooled samples of equivalent amounts of total
RNA from very young, developing, fully-opened flowers and
flowers transitioning into pods. The cDNA was
directionally ligated into the Uni-Zap XR vector
(Stratagene) and packaged using the Gigapack III Gold
packaging extracts. Phagemids containing cDNA inserts were
in vivo excised from the recombinant Uni-Zap XR vector
using ExAssist helper phage and the E. coli strain
XL1-Blue MRF' (Stratagene). Excised plasmids were plated
using SOLR cells."
```

Alignment Scores:

Pred. No.:	8.71e-93	Length:	696
Score:	831.00	Matches:	159
Percent Similarity:	86.47%	Conservative:	20
Best Local Similarity:	76.81%	Mismatches:	27
Query Match:	62.06%	Indels:	1
DB:	12	Gaps:	0

US-10-624-061-16 (1-261) x BI273174 (1-696)

QY 52 TyralaGlnleGlnGlyPheCysGlySerLeuValGluPheValVallysAsnThrGlu 71  
Db 5 TATCTTCAATCCCTGGTCTTTTGGTCTCATCTGGTTGAGTTTGTTCACCAATCTGAG 64  
QY 72 AlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGlnAsnAlaPheProGlu 91  
Db 65 ATTGTTCAAGCTAAGGTGGGAAGATACAAAACCCCTGAAGAGATCCCTTCTTCCAGAA 124  
QY 92 AsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisProGlyAla 111  
Db 125 AATTATCATGTGCTACTGTTCATCTACCCCTTCAAGAATTTTGGATCTCTGGAGCC 184  
QY 112 AlaSerIleAsnLysSerIleTrpLysVetTyrPheLysGluLeuProLeu 131  
Db 185 GCCTTCGATTAACTAAACAAGTCCTTATGGAATGATCTTTCATGACCTGCTTCATTG 244

QY 132 LeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerIleu 151  
Db 245 TTCGTTGCTTCGGTGATGATGGCAAACTATCAAACTGAGGAGCTCTCATTTG 304  
QY 152 LeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLysPheArg 171  
Db 305 TTGCAGGCCATTTCTAAAGAGTTCATTATGAAAAGTTTGTAGCTGAAGTGAATTCAGG 364  
QY 172 AspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLys 191  
Db 365 GAATCTCTCAAGACTATGAGCCTTTAATTCGGTCTAAGGACAAAGAGGGTTGATGAAA 424  
QY 192 LeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLysAlaVal 211  
Db 425 TTGTTGACATTTAAGAGTGTTCAGAGATGCTTAAGAAAAGGGTTGAAAAGAGGCAACA 484  
QY 212 ValPheGlyGlnGluValAsnLeuAsnSerAspAsnAspAsnGluAsnArgLysPhe 231  
Db 485 TTGTTTGGGCAGGATATAAGCCTTAACAGCAGTGTATGACAGTAAAGGTAAACAAGATT 544  
QY 232 AspProSerValAlaSerSerLeuTyrLysAsnTrpValIleProLeuThr-LysGluVa 251  
Db 545 GATCCATCAGTGGCTTCTAAGTTGTATGANAATGGGTATACCTCTTACTAANGANGGT 604  
QY 251 lclnValGluTyrLeuLeu 257  
Db 605 CAAGGTTGAGTACCTACTG 623

RESULT 6  
BF649665  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BF649665 605 bp mRNA linear EST 20-DEC-2000  
NF081H11EC1F1094 Elicited cell culture Medicago truncatula cDNA  
clone NF081H11EC 5', mRNA sequence.

BF649665  
BF649665.1 GI:11914795  
ESI:  
Medicago truncatula (barrel medic)  
Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.  
1 (bases 1 to 605)  
Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,  
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation -  
Center for Medicago Genomics Research  
Unpublished (2000)  
Contact: Dixon RA  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7302  
Fax: 580 221 7380  
Email: radixon@noble.org  
Insert Length: 605 Std Error: 0.00  
Plate: 081 row: H column: 11  
Seq primer: TCACACAGGAACAGCTATGAC.  
Location/Qualifiers  
1..605  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/db\_xref="taxon:3880"  
/tissue\_type="Cell cultures derived from root tissues"  
/dev\_stage="Cell suspensions were subcultured every 14  
days. Cells were induced six days after subculture"  
/clone\_lib="Elicited cell culture"  
/notes="Vector: Lambda Zap; Cells were induced with yeast  
cell wall extracts equivalent to 50ug/ml glucose in the  
final concentration. Samples were taken at 0.5, 1, 12 and  
24 hours after induction. Equal amounts of RNA from each  
time point were pooled and used for mRNA isolation."

## ORIGIN

Alignment Scores: 2,92e-88 Length: 605  
 Pred. No.: 794.00 Matches: 150  
 Score: 87.50% Conservative: 18  
 Percent Similarity: 78.12% Mismatches: 24  
 Best Local Similarity: 59.30% Indels: 0  
 Query Match: 10  
 DB: 10

US-10-624-061-16 (1-261) x BF649665 (1-605)

QY 13 AnValTyThrLeuAlaSerValArgGluAspLeuValArgGlnGluAspThrIleLe 32  
 Db 29 AATGAATATCTCTGATACAGTGGAGAGATTGGTTAGACAGAGATCAATTTGT 88  
 QY 33 TyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyAspGluLysTyr 52  
 Db 89 TTTGGTCTTATTGAGAGATCAAGTTCTCTTTAATCTCAAACTTATGATCAGAAATTAT 148  
 QY 53 AlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsnThrGluAla 72  
 Db 149 CTTCAATCCCTGGTTTGTGGCTCATTTGGTGTGAGTTTGTTCACCAATACTGAGATT 208  
 QY 73 IleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsn 92  
 Db 209 GTTCAAGCTTAAGCTGGAGATACAAACCTCTGAGAGATCCCTTCTTCCAGAAAT 268  
 QY 93 LeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAla 112  
 Db 269 TTATCATTTCTACTTGTTCATCTTACCCTTCACAAAGTTTTCATCTCGAGCGCT 328  
 QY 113 SerIleAsnIleAsnLysSerIleThrPheLysMetTyrPheLysGluLeuProLeuLeu 132  
 Db 329 TCGATTAAATAAACAGCTTATGGAAATGTACTTTCATGACCTCTCCATTTGTC 388  
 QY 133 AlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeu 152  
 Db 389 GTTGCTTCGGTGTGATGGCAACTATGCACAACTGCAGCTAGGACCTCTCATTTGTG 448  
 QY 153 GlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLysPheArgAsp 172  
 Db 449 CAGGCAATTTCTAAAGAGTTTCATTATGAAAGTTGTAGCTGAAGTGAATTCAGGAA 508  
 QY 173 AlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeu 192  
 Db 509 TCTCCTCAAGACTATGAGCTTTAATTCGGTCTAAGGACAAAGAAAGTTGATGAATTG 568  
 QY 193 LeuThrPheThrSerValGluGluThrValArgLys 204  
 Db 569 TTGACATTTAAGAGTGTGTAANAGATGTGTAAGAA 604

## RESULT 7

BF009709 620 bp mRNA linear EST 06-DEC-2001  
 LOCUS ss83b10.y1 Gm-cl064 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 DEFINITION Gm-cl064-692 5' similar to TR:Q9S7H4 Q9S7H4 CHORISMATE MUTASE ;  
 mRNA sequence.

ACCESSION BF009709.1 GI:10709985

VERSION BF009709

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

1 (bases 1 to 620)

## REFERENCE

Shoemaker R., Keim P., Vodkin L., Erpelting J., Coryell V.,  
 Khanna A., Bolla B., Marra M., Hillier L., Kucaba T., Martin J.,  
 Beck C., Wylie T., Underwood K., Steptoe M., Theising B., Allen M.,  
 Bowers Y., Person B., Swaller T., Gibbons M., Pape D., Harvey N.,  
 Schurk R., Ritter E., Kohn S., Shin T., Jackson Y., Cardenas M.,

McCann, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

High quality sequence stop: 404.

## FEATURES

Location/Qualifiers

1..620

/organism="Glycine max"

/mol\_type="mRNA"

/cultivar="Williams"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl064-692"

/tissue\_type="seedling epicotyls"

/dev\_stage="2 week old"

/lab\_host="DH10B"

/clone\_lib="Gm-cl064"

/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The cDNA library was constructed from mRNA isolated

from the epicotyls of 2 week old seedling for the cultivar

Williams. The seedlings were germinated in a growth

chamber, excised above the soil level, and the plants

were placed in a 100 ppm solution of auxin for 24 hours

prior to harvesting. Complementary DNA was synthesized

from mRNA using a primer consisting of a poly(GT)

sequence with a XhoI restriction site. EcoRI adapters

were ligated to the blunt-ended cDNA fragments followed

by XhoI digestion. The cDNA fragments were directionally

cloned into the EcoRI-XhoI restriction site of the

pluescript vector. The ligated cDNA fragments were

transformed into DH10B host cells (GibcoBRL). This

library was constructed in the laboratory of Dr. Randy

Shoemaker."

Alignment Scores: 2,23e-83 Length: 620

Pred. No.: 755.00 Matches: 156

Score: 89.13% Conservative: 8

Percent Similarity: 84.78% Mismatches: 19

Best Local Similarity: 56.39% Indels: 2

Query Match: 10 Gaps: 0

DB: 10

US-10-624-061-16 (1-261) x BF009709 (1-620)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValThrLeuAlaSerVal 20

Db 72 ATGCCCAAGACAGACAGAAAGTCTGATCTGGAAATGTACACCGTAGCTTCAGTG 131

QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLys 40

Db 132 AGGGAGGATTTGGTTAGGCAAGAGATACCATTTATGGTCTCATTCAGAGAGCCCAAG 191

QY 41 PheProSerAsnSerHisThrTyAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60

Db 192 TTCCTTAGCAATCTCTCACACCTATGATGAAGATGATGTCTCAATCCAGGTTTGTGGC 251

QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80

Db 252 TCATTGGTAGAATTGTTGTTAAGAAATACAGAGCCCATCAAGCTAAGCTGGAAGATAC 311

QY 81 LysAsnProGluGluAsnAlaPhePheProGluLysAsnLeuProProSerIleValProSer 100

Db 312 AAAAACCCCTGAAGAAAGCCCTTCTTCCAGAAAAATTTACCAACCATCAATTTGTGCCATCT 371

## ORIGIN

Alignment Scores:

Pred. No.: 2,23e-83 Length: 620

Score: 755.00 Matches: 156

Percent Similarity: 89.13% Conservative: 8

Best Local Similarity: 84.78% Mismatches: 19

Query Match: 56.39% Indels: 2

DB: 10 Gaps: 0

US-10-624-061-16 (1-261) x BF009709 (1-620)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValThrLeuAlaSerVal 20

Db 72 ATGCCCAAGACAGACAGAAAGTCTGATCTGGAAATGTACACCGTAGCTTCAGTG 131

QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLys 40

Db 132 AGGGAGGATTTGGTTAGGCAAGAGATACCATTTATGGTCTCATTCAGAGAGCCCAAG 191

QY 41 PheProSerAsnSerHisThrTyAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60

Db 192 TTCCTTAGCAATCTCTCACACCTATGATGAAGATGATGTCTCAATCCAGGTTTGTGGC 251

QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80

Db 252 TCATTGGTAGAATTGTTGTTAAGAAATACAGAGCCCATCAAGCTAAGCTGGAAGATAC 311

QY 81 LysAsnProGluGluAsnAlaPhePheProGluLysAsnLeuProProSerIleValProSer 100

Db 312 AAAAACCCCTGAAGAAAGCCCTTCTTCCAGAAAAATTTACCAACCATCAATTTGTGCCATCT 371

```

QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120
Db 372 TACTCTCTTCAACAGTTTTTGGATCTCTGGAGCTGCTTCAATTAAATTAACAGTCCATC 431
QY 121 TrpLysMetTyrPheLysGlnLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140
Db 432 TGGAAAT-TATTTCCAGAGTTACTTTTCATTTGGTTGCTTCTCGGGGATGAATCGAAC 490
QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArg-ArgIleHi 160
Db 491 TATGCCACCAACAGCTATGATCTTTTCATTTATGGAGGCGATTTTATGGAGAGATTC 550
QY 160 sTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProIle 180
Db 551 CTTGGAAGAGTTTGTACCTGAGTGAAATCAAGGATTCCTCTCCAAACTACGAAGCCTC 610
QY 180 urleArgAla 183
Db 611 TATTCGAGCT 620

RESULT 8
BI420527 605 bp mRNA linear EST 15-AUG-2001
LOCUS LjN85758b9r Lotus japonicus node library 5 and 7 week-old Lotus
DEFINITION corniculatus var. japonicus cDNA 5', mRNA sequence.
ACCESSION BI420527
VERSION BI420527.1 GI:15191550
KEYWORDS EST.
SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
REFERENCE 1 (bases 1 to 605)
AUTHORS Colebatch,G., Freund,S., Trevasakis,B and Udvardi,M.
TITLE Lotus japonicus root node ESTs: tools for functional genomics
JOURNAL Unpublished (2000)
COMMENT Contact: Udvardi MK
Molecular Plant Nutrition
Max Planck Institute of Molecular Plant Physiology
Am Muehlenberg 1, 14476 Golm, Germany
Fax: 49 331 567 8250
Email: udvardi@mpimp-golm.mpg.de
Seq primer: T7
High quality sequence stop: 605.
FEATURES
Location/Qualifiers
1..605
/organism="Lotus corniculatus var. japonicus"
/mol_type="mRNA"
/cultivar="Gifu (B-129)"
/db_xref="taxon:34305"
/dev_stage="5 and 7 week-old plants"
/clone_lib="Lotus japonicus node library 5 and 7
week-old"
/note="Organ: Nodule; Vector: pSPORT1; Site 1: SalI;
Site 2: NotI. The library was prepared using mRNA
extracted from nodules of 5 and 7 week-old Lotus plants.
Nodules were induced by, and contained Mesorhizobium
strain R7A."

ORIGIN
Alignment Scores: 1.84e-79 Length: 605
Pred. No.: 723.50 Matches: 139
Score: 86.74% Conservative: 18
Percent Similarity: 76.80% Mismatches: 19
Best Local Similarity: 54.03% Indels: 5
Query Match: 12 Gaps: 1
DB:
US-10-624-061-16 (1-261) x BI420527 (1-605)
QY 1 MetAlaLysAlaAlaGluInSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20

```

```

Db 76 ATGCCCAAGCA-----GAATCGAATGATGTGTTTCCCTGGAGTCGTG 120
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArgAlaLys 40
Db 121 AGGAGATCTGGTTAGCAAGAGATACCATTTGTTTGGTCTCATTTGACAGGCTAGG 180
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60
Db 181 TTCCCATGAATATATCACACCTATGATGATAAACTACTGGAAAATCCCGGGTTTTTGTGC 240
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80
Db 241 TCATTTGGTCGAATGTTGTTTCAAAACACATGAGGCCATTCAGCTATGGCTGGAGGTAC 300
QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100
Db 301 CAAACCCCTGAAGAAATGCTTCTTCCAGAACATTTACCATCACCAATTTGTGCCATCT 360
QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120
Db 361 TACCCCTTCACACAGATTTTGCATCTCGAGCTCTTCCATTAACATAACAGTCGATT 420
QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140
Db 421 TGGAAATTAATTTGATGATGCTTCCAAATATTGTTGCTCGCTCGGTGATGATGCAAC 480
QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHis 160
Db 481 TATGCACAACTGCAGCTAGCGATCTTTCATTTATTCAGGCGCAATCTTAGAAGATTCAT 540
QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180
Db 541 TATGGAAGTTTGTAGCTGAGCGGAATTCAGGGAATCTCTCAAGACTATGAGCCTTTA 600
QY 181 Ile 181
Db 601 ATT 603

RESULT 9
AI973566 477 bp mRNA linear EST 30-NOV-2001
LOCUS sc88e02.y1 Gm-cl019 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl019-435 5', similar to SW:CHMU_ARATH P42738 CHORISMATE MUTASE
PRECURSOR ;, mRNA sequence.
ACCESSION AI973566
VERSION AI973566.1 GI:5770392
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 477)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptos,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com

```

Insert Length: 653 Std Error: 0.00  
Seq primer: -40RP from Gibco  
High quality sequence stop: 447.  
Location/Qualifiers

## FEATURES

source

1..477  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl019-435"  
/tissue\_type="Immature seed coats of greenhouse grown plants"  
/lab\_host="DH10B (Gibco BRL)"  
/clone\_lib="Gm-cl019"  
/note="Vector: pSPORT1 (Life Technologies); Site 1: Not I; Site 2: Sal I; This cDNA library was constructed from mRNA isolated from immature seed coats (200-300 mgs) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells (Gibco BRL). This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,628-79 Length: 477  
Score: 721.00 Matches: 141  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 53.85% Indels: 0  
DB: 9 Gaps: 0

US-10-624-061-16 (1-261) x At973566 (1-477)

QY 121 TrrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140  
Db 1 TGGAAAGTATTTCAGAGAGTTACTTCATGCTTGTCTACTTCGGGTGATGCGCAAC 60  
QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHis 160  
Db 61 TATGCGCAACTGCAGCTTAATGACCTTTTCATTATTCAGCTCCATCTCTAGAGGATTTCAC 120  
QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
Db 121 TATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCAAGACTACGAGCCCTTA 180  
QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200  
Db 181 ATTCGAGCTAAGGATAAAGAGGATTGATGAATTTGTGACATTTACCAAGCGTTGAAGAG 240  
QY 201 ThrValArgLysArgValGluLysIleAlaValValPheGlyGlnGluValAsnLeuAsn 220  
Db 241 ACGGTGAGGAAGAGATTGAAAAGAGCTGTGTGTGTTGGCGAGGAAGTGAATCTTAAC 300  
QY 221 SerAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240  
Db 301 AGTGATGACATGACATGAATAACCGTAATTTTCATCCATCAGTGGCTTCTAGCTGTAC 360  
QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrIleLeuArgArgLeu 260  
Db 361 AAAAATGGGTGATACCTCTCTCACCAGAGGTTTCAGTTGAGTACCTCTTTCGCCCGCTCA 420  
QY 261 Asp 261  
Db 421 GAC 423

## RESULT 10

BE821647/c

## LOCUS

DEFINITION  
GN70001SALO11 Gm-r1070 Glycine max cDNA clone Gm-r1070-5781 3',  
mRNA sequence.  
ACCESSION  
BE821647  
VERSION  
BE821647.1 GI:10253881  
KEYWORDS  
EST.  
SOURCE  
Glycine max (soybean)  
ORGANISM

## REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1..(bases 1 to 560)

Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Corvelli,V.,

Epeidig,J., Rapp,C., Shoop,E., Fardinas,J., Liu,D. and Lewin,H.

A Functional Genomics Program for Soybean (NSF 9872565)

Unpublished (1999)

Other ESTs: At973566 corresponding to Gm-cl019-435 (5')

Contact: Vodkin, L.O., PI, A Functional Genomics Program for

Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional

Genomics

University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA

Tel: (217) 244-6147

Fax: (217) 333-4582

Email: l-vodkin@uiuc.edu

This clone is available through: Genome Systems, Inc. 4633 World

Parkway Circle St. Louis, Missouri 63134. For further information

call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)

427-3324 or contact: clones@genomesystems.com or info@genomes

systems.com web site: www.genomesystems.com

Seq primer: 5'-TTTTTTTTTTTTTTTTTTT(A/C/G)-3'.

Location/Qualifiers

1..560

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone\_lib="Gm-r1070-5781"

/clone\_lib="Gm-r1070"

/note="The library Gm-r1070 is a sequence-driven, reracked

set of 9,216 clones selected from cDNA libraries from

various tissues and stages of development of soybean that

represent 2,639 sequences from immature cotyledons, 1,770

from immature seed coats, 3,938 from flowers, and 869

from young pods. The 5' ESTs of the source clones from

the different libraries was used to select singletons, or

a representative of each contig, which were reracked to

form library Gm-r1070. The cDNA clones of the reracked

Gm-r1070 library were then sequenced at the 3' end. The

contig analysis to select unique genes was performed by

the laboratory of Ernest Retzel, Center for Computational

Genomics and Bioinformatics, University of Minnesota,

http://www.cbc.umn.edu/researchprojects/soybean/index.html

Reracking was performed by Genome Systems, St. Louis,

http://www.genomesystems.com, and 3' sequencing by the

Keck Center for Comparative and Functional Genomics,

University of Illinois,

http://www.life.uiuc.edu/biotech/keck.html. Note: The

corresponding 5' EST from each clone in the Gm-r1070

library is listed in the 'OTHER EST' field. The detailed

information on the source library for each clone can also

be obtained by referring to the Genome Systems clone ID of

the original cDNA library that is also listed under

'OTHER EST'."

## ORIGIN

## Alignment Scores:

Pred. No.: 3,358-79 Length: 560  
Score: 721.00 Matches: 141  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 53.85% Indels: 0



```

DB:          10          Gaps:          0
US-10-624-061-16 (1-261) x BE821647 (1-560)

QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyValAsn 140
DB 547 TGGAAATGTAATTCAGAGAGTTACTTCCATTGCTTGCTTACTCGGTGATGATGCAAC 488
QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgIleHis 160
DB 487 TATGCGCAATGACGCTAATGACCTTTCATTATTGTCAGTCCATCTCTAGAGGATTCA 428
QY 161 TyrGlyLysPheValAlaGlnValLysPheArgAspAlaProGlnAspTyrGluProLeu 180
DB 427 TATGGAAGTTGTAGCTGAGGTGAATTCAGGATGCTCCCAAGACTACGAGCTTTA 368
QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200
DB 367 ATTGAGCTAAGTAAGAAGATTGATGMAATTGTTGACATTTACAGCGTTGAAGAG 308
QY 201 ThrValArgLysArgValGlnLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220
DB 307 ACGGTGAGGAAGAGAGTTGAAAAGAGGCTGTGTGTTTGGGAGGAGGAACTCTTAAC 248
QY 221 SerAspAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240
DB 247 AGTGATGACAAATGACAAATGAACCGTAATTTGATCCATGATGCTTCTAGCTTGAC 188
QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGlnValGlnValGlnValGln 260
DB 187 AAAAAATTGGTGATACCTCTCACCAAGGAGGTTGAGTTGAGTACCTCTTGGCGGCTCA 128
QY 261 Asp 261
DB 127 GAC 125

RESULT 11
CB827129
LOCUS      529 bp      mRNA      linear      EST 16-MAY-2003
DEFINITION LJN87067r Lotus japonicus nodule library 5 and 7 week-old Lotus
            corniculatus var. japonicus cDNA 5', mRNA sequence.
ACCESSION  CB827129
VERSION     1
KEYWORDS   EST.
SOURCE     Lotus
ORGANISM   Lotus corniculatus var. japonicus (Lotus japonicus)
            Lotus corniculatus var. japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
            Lotus.
REFERENCE  1 (bases 1 to 529)
            Colebatch, G., Freund, S., Trevaskis, B and Udvardi, M.
            Lotus japonicus root nodule ESTs: tools for functional genomics
            Unpublished (2000)
            Contact: Udvardi MK
            Molecular Plant Nutrition
            Max planck Institute of Molecular Plant Physiology
            Am Muehlenberg 1, 14476 Golm, Germany
            Fax: 49 331 567 8250
            Email: udvardi@mpimp-golm.mpg.de
            Seq primer: 17
            High quality sequence stop: 529.
            Location/Qualifiers
                1..529
                    /organism="Lotus corniculatus var. japonicus"
                    /mol_type="mRNA"
                    /cultiVar="Gifu (B-129)"
                    /db_xref="taxon:34305"
                    /dev_stage="5 and 7 week-old plants"
                    /clone_lib="Lotus japonicus nodule library 5 and 7
                    week-old"
                    /note="Organ: Nodule; Vector: pSPORT1; Site 1: SalI;
                    Site 2: NotI; The library was prepared using mRNA

```

extracted from nodules of 5 and 7 week-old Lotus plants.  
Nodules were induced by, and contained Mesorhizobium  
strain R7A."

## ORIGIN

```

Alignment Scores:          Length:          529
Pred. No.:                717.00          Matches:          140
Score:                    98.82%          Conservative:      11
Percent Similarity:      82.35%          Mismatches:       19
Best Local Similarity:   53.55%          Indels:           0
Query Match:             14              Gaps:             0
DB:
US-10-624-061-16 (1-261) x CB827129 (1-529)
QY 54 GlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysValLysAsnThrGluAlaIle 73
DB 5 AAAATCCCGGGTTTTTGTGGCTCATTTGTCGAATGTGTTTCAAAACACTGAGGCCATT 64
QY 74 GlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeu 93
DB 65 CAAGCTATGGCTGGAAGTACCAAAACCTGAAGAAATGCTTCTTCCAGAACATTTA 124
QY 94 ProProSerIleValProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSer 113
DB 125 CCATCCCAATTTGTGCCATCTTACCCCTTCACACAGATTTTGCATCTCGAGCTGCTTC 184
QY 114 IleAsnIleAsnLysSerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAla 133
DB 195 ATTAACATAAACAAGTCGATTGGAAAAATTTACTTGATGAGTTGCTTCCAAATATTGTT 244
QY 134 ThrSerGlyAspAspGlyValAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGln 153
DB 245 CGTCCGGTGATGATGGCAACTATGCAAACTGCAGTAGCGATCTTTCATTATTGCAG 304
QY 154 SerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLysPheArgAspAla 173
DB 305 GCAATCTCTAGAGGATTTCATTATGGAAGTTTGTAGCTGAGGCGAAATTCAGGGAATCT 364
QY 174 ProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeu 193
DB 365 CTTCAAGACTATGAGCTTTTAAATTCGTCTAAGCACACAGAGGGTTGATGAATATTG 424
QY 194 ThrPheThrSerValGluGluThrValArgLysArgValGluLysLysAlaValValPhe 213
DB 425 ACATTTGAGAGTGTTTCAAGAGATGTTGATAAAGAGAGTTGAAAAGAGGCCATGTTGTT 484
QY 214 GlyGlnGluValAsnLeuAsnSerAspAsp 223
DB 485 GNGCAGGAGGTAGGCTTAACAGTGATGAT 514

```

## RESULT 12

BG239446

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 422)

Shoemaker, R., Keim, P., Vodkin, L., Expelting, J., Corvelli, V.,

Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,

Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,

Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,

Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCaun, R., Waterston, R. and Wilson, R.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

Glycine max (soybean)

Glycine max

EST.

GI:12774519

BG239446.1

422 bp

mRNA

linear

EST 28-NOV-2001

sab70d03.y1 Unknown Library Type Glycine max cDNA clone GENOME

SYSTEMS CLONE ID: Gm-C1032-2453 5' similar to TR:Q9S7H4 Q9S7H4

CHORISMATE MUTASE ; mRNA sequence.

TITLE Public Soybean EST Project  
JOURNAL Unpublished (1999)  
COMMENT Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estowatson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
High quality sequence stop: 402.

## FEATURES

source  
1. .422  
Location/Qualifiers  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: GM-cl032-2453"  
/tissue\_type="unknown"  
/lab\_host="unknown"  
/clone\_lib="unknown Library Type"  
/note="Vector: unknown; Site 1: unknown; Site 2: unknown;  
This library is assigned to clones that are of uncertain  
origin due to an error during re-array, prepping, and/or  
analysis. Samples assigned this library are still genuine  
Glycine Max cDNA sequence, but we are unsure of its source  
and identification. Due to the clone id being  
incorrectly assigned, THESE CLONES ARE NOT AVAILABLE TO  
ORDER."

## ORIGIN

Alignment Scores:  
Pred. No.: 5,11e-77 Length: 422  
Score: 702.00 Matches: 138  
Percent Similarity: 98.57% Conservative: 0  
Best Local Similarity: 98.57% Mismatches: 2  
Query Match: 52.43% Indels: 0  
DB: 12 Gaps: 0

US-10-624-061-16 (1-261) x BG239446 (1-422)

QY 113 SerLeuSerGlyValAspGlyValSerLeuTrpLysMetTyrPheLysGluLeuLeuProLeuLeu 132  
DB 3 TCAATTAAACATAAACAAGTCCATCTGGAATATGATTTTCAAGAGTTACTTCCATTGCTT 62  
QY 133 AlaThrSerGlyValAspGlyValSerLeuTrpLysMetTyrPheLysGluLeuLeuProLeuLeu 152  
DB 63 GTTACTTGGGTGATGATGGCACTATCGCAACTGCGACTATGACCTTTTCAATTATG 122  
QY 153 GlnSerLeuSerArgArgLeuHisTyrGlyLysPheValAlaGluValLysPheArgAsp 172  
DB 123 CAGTCCATCTCTAGAAGGATTCACTATGAAAGTTGTAGCTGAGGTGAAATTCAGGAT 182  
QY 173 AlaProGlnAspTyrGluProLeuLeuArgAlaLysAspLysGluLeuMetLysLeu 192  
DB 183 GTCCTCAGACTACGAGCCCTTTAATTCGACTAAGGATAAAGAGATTGATGAATG 242  
QY 193 LeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLysAlaValVal 212  
DB 243 TTGACATTTTACAAGCCGTGAGAGAACGCTGAGGAGAGAGTTGAAAGAGGCTGTGTG 302  
QY 213 PheGlyGlnGluValAsnLysSerAspAspAsnAspAsnGluAsnArgLysPheAsp 232  
DB 303 TTGGGCGAGGAGTGAATCTTAAACAGTATGACATGACATGAAATGAAATTCAT 362  
QY 233 ProSerValAlaSerSerLeuTyrLysAsnTrpValLeuProLeuThrLysGluValGln 252  
DB 363 CCATCAGTGGCTTCTAGCTTTGTACAAAATTTGGGTGATACCTCTCACCAAGGAGTTT 422

RESULT 13

CD427426

LOCUS

810 bp mRNA linear EST 02-JUN-2003

## DEFINITION

SAL\_30\_H03\_g1\_A002 Salicylic acid-treated seedlings Sorghum bicolor  
cDNA clone SAL\_30\_H03\_A002 5', mRNA sequence.

## ACCESSION

CD427426

## VERSION

CD427426.1 GI:313333689

## KEYWORDS

EST.

## SOURCE

Sorghum bicolor (sorghum)

## ORGANISM

Sorghum bicolor

## REFERENCE

1 (bases 1 to 810)

## AUTHORS

Cordonnier-Pratt M.-M., Wentzel, V., Suzuki, Y., Sugano, S.,  
Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Salzman, R.,  
Chua Tan, N., Gonzalez, M., Lane, S., Miller, V., Nanda, P.,  
Glaseinde, O., Eastman, A. and Pratt, L.H.

## TITLE

An EST database from Sorghum: salicylic acid-treated seedlings

## JOURNAL

Unpublished (2003)

## COMMENT

Other ESTs: SAL\_30\_H03.bi\_A002

Contact: Cordonnier-Pratt M.-M.

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmp@pratt@uga.edu

Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in

the Human Genome Center, University of Tokyo Institute of Medical

Science, plant material and RNA prepared at Texas A & M University,

sequencing done in the Laboratory for Genomics and Bioinformatics,

University of Georgia. Sequence ends have been trimmed to exclude

vector and regions below Phred quality 16. Three-prime sequences

are presented as their reverse complement and have been trimmed to

exclude polyA.

Seq primer: Sug5 (CTTCTGCTCTAAAGCTGCG).

Location/Qualifiers

1. .810

/organism="Sorghum bicolor"

/mol\_type="mRNA"

/cultivar="BTx623"

/db\_xref="taxon:4558"

/clone="SAL\_30\_H03\_A002"

/lab\_host="DH10B-TI phage-resistant E. coli"

/clone\_lib="Salicylic acid-treated seedlings"

/note="Vector: pME18S-FU3; Site 1: XhoI; Site 2: XhoI; The

library was prepared from polyA+ RNA from seedlings grown

in hydroponic culture. At 8 days of age, medium was

supplemented with 1 mM salicylic acid (SA). Roots and

shoots were harvested after 27 and 72 hr and material from

both time points was combined prior to RNA isolation.

Double-stranded cDNA was cloned unidirectionally into

different DraIII sites of the pME18S-FU3 vector (5-prime

DraIII site is CACTGTGTG, 3-prime DraIII site is

CACCATGTG). XhoI excises the cDNA insert."

## ORIGIN

Alignment Scores:

Pred. No.: 8,92e-74 Length: 810

Score: 679.50 Matches: 139

Percent Similarity: 69.84% Conservative: 37

Best Local Similarity: 55.16% Mismatches: 71

Query Match: 50.75% Indels: 5

DB: 14 Gaps: 3

US-10-624-061-16 (1-261) x CD427426 (1-810)

QY 12 GlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAspThrIle 31

DB 26 GCGGAGCAGCTGAGCTGCGCGGTGCGGATGGCTGGTCCGCTCGAGACTCCGTG 85

QY 32 IleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAspGluLys 51

DB 86 GTGTTGGGCTCATCGAGCGGCCCGGATCCGCGAAGCGCGGCTACGCGGCCGCC 145

```

QY 52 TyAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValysAsnThrGlu 71
Db 146 GCGCGGGCGGGCGGAGGATAT-----TCGCTCGTGAGTTCTTCGCGGAAGCAGAG 199
QY 72 AlaIleGlnAlaIlyslaglyArgTyrLysAsnProGluGluAsnAlaPheProGlu 91
Db 200 GCGCTTCAACGCGAAGGCTGCATATCAAAAGCCAGAGATGTTCCATCTTCCTCAA 259
QY 92 AsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisProGlyAla 111
Db 260 GATCTTCCCTCGGCTCTCTTCTACCAAGCCTTCCCAAGGCTTTGCACCTTTTGT 319
QY 112 AlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeuProLeu 131
Db 320 TTGTTGCTCACTGTAATGATGCAATATGAAATATATTTTTCATGAATGTCCTCATTA 379
QY 132 LeuAlaThrSerGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeu 151
Db 380 TTCACCTGTTGATGCTGATGATGTAATGATGATGATGATGATGATGATGATGATG 439
QY 152 LeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLysPheArg 171
Db 440 CTCAGGTTCTATCAAAAGATTCATATTTGGTAAATATGTTGCTGAGGTCAAGTTCAA 499
QY 172 AspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLys 191
Db 500 GATGCTCTCAAGATTATGCTGATTAATAAAGCAAGACAGCAATCTTTTGATGAT 559
QY 192 LeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLysAlaVal 211
Db 560 CTGCTGACATTTCAAGCTGTGTGAAGAGAGAGTTTAAAGAGAGATGAGAGAAGGCCAG 619
QY 212 ValPheGlyGlnGluValAsnLeu-----AsnSerAspAspAsnAspAsnGluAsnArg 229
Db 620 ACCTTCGGCGAGAACCTCACTTCGGAGGACAATGCCACTGCCAGTCACAGCGAA---TGC 676
QY 230 LysPheAspProSerValAlaSerSerLeuTyrLysAsnTyrValIleProLeuThrLys 249
Db 677 AAGTCAATCTTAAGTGCTTCCAGCTGATGATGATGATGATGATGATGATGATGATG 736
QY 250 GluValGlnValGluTyrLeuLeuArgArgLeuAsp 261
Db 737 GAGTCGAGTCGAGTATCTCTCGCGCGCTCGAC 772

RESULT 14
LOCUS BM093460
DEFINITION saJ09c02.y1 Gm-c1065 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-c1065-9724 5', similar to TR:Q957H4 Q957H4 CHORISMATE MUTASE
/, mRNA sequence.
ACCESSION BM093460
VERSION BM093460.1 GI:17022426
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max (soybean)
REFERENCE 1 (bases 1 to 560)
AUTHORS Shoemaker,R., Reim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
PUBLIC SOYBEAN EST PROJECT
JOURNAL Public Soybean EST Project
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

```

## FEATURES

source

1. 560

```

/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1065-9724"
/tissue_type="germinating shoots"
/lab_host="DH10B"
/clone_lib="Gm-c1065"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
germinating shoots of the cultivar Williams. The seeds
were allowed to germinate for 24 hours prior to being
cold stressed for 2 days at 4C. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy
Shoemaker."

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 4,38e-73 Length: 560
Score: 672.00 Matches: 140
Percent Similarity: 84.85% Conservative: 0
Best Local Similarity: 84.85% Mismatches: 25
Query Match: 50.19% Indels: 25
DB: 12 Gaps: 1
US-10-624-061-16 (1-261) x BM093460 (1-560)

```

```

QY 122 LysMetTyrPheLysGluLeuProLeuLeuAlaThrSerGlyAspAspGlyAsnTyr 141
Db 3 AAAATGATATTTCAAGAGATTTACTTCCATTTGCTTCTTCTGGGTGATGATGGCACTAT 62
QY 142 AlaGlnThrAlaAlaAsnAspLeuSerLeu----- 151
Db 63 GCGCAACTGCGCTAATGACCTTTTCATTATTCAGGACCACCACTCAAGATTTTGTAGTC 122
QY 152 -----LeuGlnSerIleSerAr 157
Db 123 TCCTAAACTAGTATTGATCTCTGACTGATCATTTCTATAAAACATTTGCAGTCCATCTCTAG 182
QY 157 GAGTleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTy 177
Db 183 AAGGATTCACATGATGAAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCTCAAGACTA 242
QY 177 rGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSe 197
Db 243 CGAGCTTTTAAATTCAGCTAAGGATAAAGAGGATTGATGAATTTGTTGACATTTACAAG 302
QY 197 rValGluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluVa 217
Db 303 CGTTGAAGAGACGGTGAAGAGAGAGATTTGAAAGAAAGGCTGTGTGTTTGGCGAGGAAGT 362
QY 217 LasnLeuAsnSerAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaLase 237
Db 363 GATCTTTACAGTGTATGACATGACATGAAACCGTAAATTTGATCCATCAGTGGCTTC 422
QY 237 rSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLe 257
Db 423 TAGCTTTGTACAAAAATTTGGGTGATACCTCTCACCAAGAGGTTTGTAGTGTAGTACCTCTT 482

```

